PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 13 November 15, 1927 Number 11

A PERIODIC TABLE OF THE STRUCTURE OF ATOMS AND ITS RELATION TO ION FORMATION AND VALENCE

By Arthur A. Noyes and Arnold O. Beckman

Gates Chemical Laboratory, California Institute of Technology

Communicated September 14, 1927

Introduction.—For the purpose of studying the interpretations which the recently acquired knowledge of the structure of atoms might afford as to tendencies of elements to form ions and to exist in definite valence states, the authors compiled from spectroscopic publications of the last three years the values of the energies attending the removal of the various electrons from the outer shells of the atoms and ions of the first twenty-six elements (higher elements not being included, since the spectral data have as yet been too incompletely interpreted). And a chart was prepared showing the number, the quantum states, and the energies of removal of these electrons. The data upon which the chart is based and the detailed discussion of it in relation to the chemical properties mentioned are to be published in full in the *Journal of the American Chemical Society;** but the interest which the matter might have for scientists in other fields has led to the presentation in these Proceedings of the chart and a summary of the conclusions reached as to ion formation and valence.

In the assembling and interpretation of the spectroscopic data we have been greatly assisted by the advice of Prof. Ira S. Bowen of this Institute, by whose experimental researches, made in association with Dr. R. A. Millikan, the data for the chart have in large measure been made available. This paper has, moreover, been possible only through the recent interpretations of electron distribution by Main-Smith[1] and Stoner[2] and of complex spectra by Russell and Saunders,[3] Pauli,[4] Heisenberg[5] and Hund.[6]

Description of the Chart.—The chart is shown on page 740. Along the axis of abscissas is plotted the atomic number of the element, which also represents the positive charge on the nucleus of its atom and ions, as well as the total number of electrons around the nucleus of the

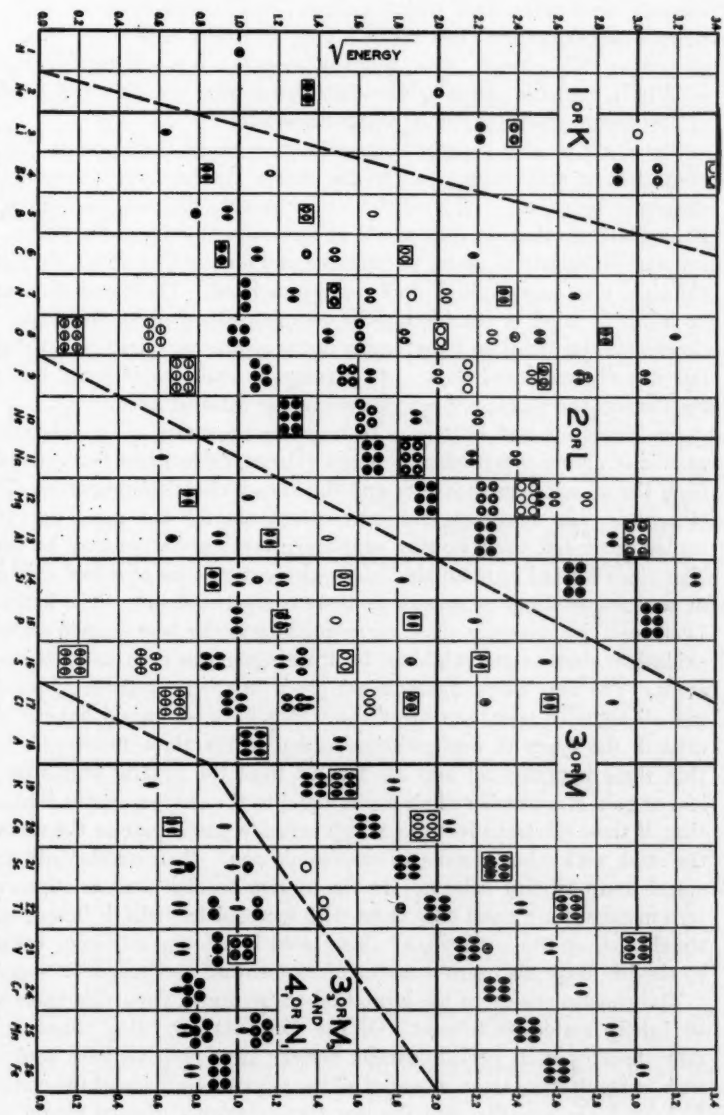
* The chart was originally shown and discussed in an address before the Pacific Division of the American Association for the Advancement of Science at its Reno meeting on June 22, 1927.

neutral atom. As ordinates are plotted the square-roots of the energies absorbed in the removal of the respective electrons from the atom in its normal state in terms of the energy asborbed in removing the electron from the hydrogen atom as unity; in other words, the square-roots of the ratios ν/R of the wave-number of the spectral term to the Rydberg number. The square-roots are plotted, rather than the energies themselves, since this enables higher energy values to be included in the chart without unduly extending it, and since Millikan and Bowen[7] have found that Moseley's principle that the square-roots of the energies are nearly linear functions of the atomic numbers applies to the corresponding electrons of ions that have the same configuration but different nuclear charges.

The number of electrons having a definite energy of removal is shown by a corresponding number of circles or ellipses; it being understood that the exact value of the energy is at the center around which these electrons are grouped. Black circles or ellipses are used for electrons of the neutral atom; black circles or ellipses with a white circle at the center for the unipositive ions; white circles or ellipses for bipositive ions; white circles or ellipses with a numeral (3, 4 or 5) within for multipositive ions; and white circles with one or two cross bars for uninegative or binegative ions.

The quantum states of the electrons are differentiated by the use of circles and of ellipses of different eccentricity. Namely, in correspondence with the Bohr orbital theory, when the principal and azimuthal quantum numbers (n and k) are equal (as in the 1_1, 2_2, 3_3 ... quantum states) the electrons are represented by circles; and when these quantum numbers are unequal, ellipses are used whose major and minor axes have the same ratio as the two quantum numbers, thus 2:1 for the 2_1 state; 3:2 for the 3_2 state; etc. The principal quantum numbers (1, 2, 3, 4) or the shells (K, L, M, N) of the electrons are also indicated by the regions of the plot in which they lie, these regions being separated from each other by heavy oblique broken lines. It is not possible, however, to separate completely in this way the 3 or M from the 4 or N electrons, for the electrons of the 3_3 and 4_1 quantum types have energies of removal of the same order of magnitude. In the lower right-hand region, therefore, the circles are to be understood to represent 3_3 electrons and the ellipses 4_1 electrons.

On the chart are also differentiated the atom configurations for which the inner quantum number J is zero and the magnetic moment μ is zero from those for which these quantities have finite values; namely, by enclosing in the former case (where $J = 0$ and $\mu = 0$) the lowest group of electrons within a rectangle. It may be recalled that the quantity J is an integer or half-integer expressing the quantized resultant, for any given atom or ion, of the moments of momentum (in terms of $h/2\pi$) of the orbital rotations and spinnings about their axes of all the individual electrons, and that μ is the corresponding resultant for the magnetic


√ENERGY
1 OR K
2 OR L
3 OR M
3 OR M AND 4 OR N
●.. ELECTRONS OF NEUTRAL ATOM;
●.. OF UNIPOSITIVE ION;
O●.. OF BIPOSITIVE ION;
⊙●.. OF TRIPOSITIVE ION;
⊙ OR ⊙.. OF UNI- OR BINEGATIVE ION.

moments of the electrons. The values of J for all the neutral atoms have recently been determined by McLennan, McLay and Smith[8] for all the neutral atoms in their most stable states, and their results are those employed in this article. The values of μ, which is equal to the product of J by a variable factor g, were calculated with the aid of the values of this factor given by Hund[9] and Sommerfeld.[10]

The chart gives a complete representation of all the electrons for the neutral atom and its ions only in the case of the first four elements (hydrogen to beryllium). For all the other elements the two electrons in the K shell are omitted, as they would lie far above the top of the chart; and for the neutral atoms above phosphorus and for the ions above aluminum the eight electrons of the L shell are also omitted. These omissions must be borne in mind in considering the complete structure of the atom. It should also be mentioned that many of the points are based on estimated, not directly measured data. No attempt is made to indicate this distinction on the chart; but it is shown in the detailed article.

Ion-Formation and Valence.—To facilitate the study of the relations of atomic structure to ion-formation and valence, table 1 has been prepared from the atomic-structure data and the known chemical properties of the elements. The headings and conventions are for the most part self-explanatory; but the following supplementary statements may be made. The "ions formed" are all the simple elementary ions of whose existence in aqueous solutions or in crystals there is good evidence. The indicated "valences" include only those corresponding to the best defined states of oxidation, such as are exhibited by the elements in their salts, bases or acids. The ions whose symbols are given in the next three columns do not all actually exist, but are those which have, or would have if they existed, the types of configuration named in the three headings. Note that these hypothetical ions are formed from the neutral atoms by the loss or gain of a number of electrons equal to the charges on the ions, and that, if these electrons be those which actually participate in the union of the atom with other atoms (*the reacting electrons*), their number will correspond to an existing valence. In the column headed "ions or atoms with zero moment" are given only those with incomplete shells, it being understood that, as above stated, all those cited in the two adjoining columns to the left have also zero moment of momentum and magnetic moment.

The conclusions as to ion-formation to be drawn from this table may be briefly stated as follows. Of the elements preceding titanium the first three of each period (except boron) form positive ions, and only such as result from the removal of all the electrons outside of the complete 1, 2, or $3_1 + 3_2$ shells; and the last three elements of each period alone form negative ions, and only such as result from the completion of the 2 or $3_1 + 3_2$ shells with electrons. Intermediate ions, such as Al^+ or Ca^+,

TABLE 1

ION FORMATION AND VALENCE IN RELATION TO ATOMIC STRUCTURE

Element	Atomic no.	Ions formed	Valences	Ions or Atoms with Complete Shell	Ions or Atoms with Complete Shell	Ions or Atoms with Zero moment	M	Energy of removal of outer electrons[11] from M^+	M^{2+}	M^{3+}	M^{4+}
H	1	H^+, H^-	1 −1	H^+, H^-	..	...	1.00	∞	...	...	...
He	2	None	0	He	..	...	1.80	4.00	∞	...	...
Li	3	Li^+	1	Li^+	..	...	0.40	5.6	9.00	∞	...
Be	4	Be^{2+}	2	Be^{2+}	..	...	0.70	1.34	11.4	16.0	∞
B	5	...	3	B^{3+}	B^+	...	0.62	1.79	2.80	19.3	25.0
C	6	...	4	C^{4+}	C^{2+}	...	0.83	1.79	3.36	4.74	29.0
N	7	...	3, 5 −3	N^{5+} N^{3-}	N^{3+}	N^+	1.07	2.18	3.49	5.43	7.19
O	8	... O^{2-}	... −2	O^{6+} O^{2-}	O^{4+}	O^{2+}	1.00	2.58	4.05	5.69	8.07
F	9	F^-	−1	F^{7+} F^-	F^{5+}	F^{3+}	1.25	2.39	4.62	6.40	7.56
Ne	10	None	0	Ne	..	...	1.59	2.70	...	...	...
Na	11	Na^+	1	Na^+	..	...	0.38	3.50	...	...	...
Mg	12	Mg^{2+}	2	Mg^{2+}	..	...	0.56	1.11	5.98	...	...
Al	13	Al^{3+}	3	Al^{3+}	Al^+	...	0.44	1.34	2.09	9.00	...
Si	14	...	4	Si^{4+}	Si^{2+}	...	0.76	1.20	2.34	3.32	12.47
P	15	...	1, 3, 5 −3	P^{5+} P^{3-}	P^{3+}	P^+	0.98	1.46	2.22	3.53	4.78
S	16	... S^{2-}	4, 6, −2	S^{6+} S^{2-}	S^{4+}	S^{2+}	0.76	1.74	2.37	3.48	4.93
Cl	17	... Cl^-	1, 3, 5, 7 −1	Cl^{7+} Cl^-	Cl^{5+}	Cl^{3+}	0.96	1.66	2.76	3.50	5.00
A	18	None	0	A	..	...	1.16	..	...	...	...
K	19	K^+	1	K^+	..	...	0.32	2.34	...	...	...
Ca	20	Ca^{2+}	2	Ca^{2+}	..	...	0.45	0.87	3.76	...	...
Sc	21	Sc^{3+}	3	Sc^{3+}	..	...	0.49	0.95	1.82	5.33	...
Ti	22	Ti^{2+} Ti^{3+} Ti^{4+}	2, 3, 4	Ti^{4+}	..	...	0.50	1.00	2.04	3.18	7.07
V	23	V^{2+} V^{3+}	2, 3, 4, 5	V^{5+}	..	V^+	0.52	1.04	...	...	5.08
Cr	24	Cr^{2+} Cr^{3+}	2, 3, 6	Cr^{6+}	..	Cr^{3+}	0.50	..	...	...	...
Mn	25	Mn^{2+} Mn^{3+} Mn^{4+}	2, 3, 4, 6, 7	Mn^{7+}	..	Mn^{3+}	0.55	1.07	...	...	...
Fe	26	Fe^{2+} Fe^{3+}	2, 3, 6	Fe^{6+}	..	Fe^{4+}	0.59	..	...	...	...

do not form as they do at high temperatures, even though the energy required for the removal of the first electron is much smaller than that for the second and even when, as with Al^+, the resulting ion would have its lowest energy electrons in a complete subshell (3_1), which has, moreover, zero moment of momentum and zero magnetic moment. Titanium, vanadium, chromium, manganese and iron all form bipositive ions through the loss of their 4_1 electrons, even though the remaining electrons do not suffice to complete the 3_3 shell; the incomplete 3_3 shell being apparently as stable as the complete 2 quantum shell. They also form other ions which have no apparent relation to the structure of their atoms.

The conclusions in regard to valence may be summarized as follows. None of the first four elements of the first two periods and neither of the first two elements of the third period exhibit any other well-defined valence than that which would result if the electrons giving rise to the valence were those, and only those outside of the complete 1, 2 or $3_1 + 3_2$ quantum shells. The possibility that, in the cases of boron and aluminum or of carbon and silicon, the electrons outside of a complete subshell (2_1 or 3_1) (with its zero moment of momentum and zero magnetic moment) might alone function as valence electrons does not cause these elements to exist in the univalent or bivalent state.

The elements nitrogen, phosphorus, sulphur and chlorine, on the other hand, exhibit the positive valences that would result if the reacting electrons are those outside: (1) the complete shells 1 or 2; (2) the complete subshells 2_1 or 3_1; and (3) the configurations other than these that have zero moment of momentum and zero magnetic moment. These elements also exhibit negative valences of the type (1), namely, those that correspond to the taking up of enough more electrons to complete the 2 or $3_1 + 3_2$ shells. Thus, the valences corresponding to the three types are:

(1) Nitrogen and phosphorus, +5 and −3; sulphur, +6 and −2; chlorine, +7 and −1.

(2) Nitrogen, +3; phosphorus, +3; sulphur, +4; chlorine, +5.

(3) Nitrogen, +1; phosphorus, +1; sulphur, +2; chlorine, +3.

And, in fact, these elements exist in the form of salts in all these valences, with the exception of unipositive nitrogen (which, however, forms N_2O) and possibly of bipositive sulphur (which, however, forms derivatives of the acid H_2SO_2). Furthermore, the chemical stability of these various valences is clearly related to the type, being greatest for type (1) and least for type (3). Thus, type (1) is that of the very stable nitrates, phosphates, sulphates and perchlorates, as well as of the amines, phosphines, sulphides and chlorides; type (2) is that of the much less stable nitrites, phosphites, sulphites and chlorates; and type (3) is that of the unstable hypophosphites and chlorites and of the non-existent or extremely unstable salts M^INO

and $M^{I}_{2}SO_2$. These configuration types do not account for unipositive chlorine (in hypochlorites); but this is perhaps related to its instability; nor do they explain the existence of isolated compounds with even valences, such as NO, NO_2, ClO_2. The three types together account for the long-known fact that the well-defined valences are all even (as in the case of sulphur) or all odd (as in the case of nitrogen, phosphorus and chlorine).

The six elements beyond calcium in the third period exhibit positive valences that correspond to the complete $3_1 + 3_2$ shell. Moreover, all of them (except scandium) form bivalent compounds that can be accounted for by assuming that only the two 4_1 electrons are reacting, which implies that the underlying 3_2 shell is an exceptionally stable configuration even when it is incompletely filled with electrons. The other irregular valences exhibited by these elements cannot be explained by our present knowledge of the structure of their atoms. These valences evidently constitute a phenomenon of a type distinct from the regular even or odd valences of the nitrogen, sulphur and chlorine groups.

In general, the energies of removal of the successive electrons seem to play only a secondary part in determining the valences that exist.

[1] Main-Smith, *Chemistry and Atomic Structure*, D. Van Nostrand Co., New York, 1924, pp. 189–198.

[2] Stoner, *Phil. Mag.*, **48**, 719 (1924).

[3] Russell and Saunders, *Astrophys. J.*, **61**, 38 (1925).

[4] Pauli, *Zs. Physik*, **31**, 765 (1925).

[5] Heisenberg, *Ibid.*, **32**, 841 (1925).

[6] Hund, *Ibid.*, **33**, 345 (1925); **34**, 296 (1925).

[7] Millikan and Bowen, these PROCEEDINGS, **13**, 531 (1927).

[8] McLennan, McLay and Smith, *Proc. Roy. Soc.*, **A112**, 76 (1926).

[9] Hund, *Linienspectren und Periodisches System der Elemente*, 1927, p. 106.

[10] Sommerfeld, *Atombau und Spectrallinien*, 4th edition, 1924, p. 623.

[11] The energies of removal of the outer electrons from negative ions are: O^-, 0.34; O^{2-}, 0.03; F^-, 0.52; S^-, 0.30; S^{2-}, 0.03; Cl^-, 0.45. All these values were obtained by linear extrapolation on the square-root plot.

THE MAGNETIC BEHAVIOR OF OZONE

By OLIVER R. WULF*

CHEMICAL LABORATORY, UNIVERSITY OF CALIFORNIA

Communicated October 14, 1927

Professor G. N. Lewis suggested to the author that it would be interesting to measure the magnetic susceptibility of ozone. Becquerel[1] made such a measurement in 1881 and concluded that ozone was considerably more paramagnetic than oxygen. In the same year Schumeister,[2] while studying the magnetic behavior of oxygen, was led to a similar conclusion. Professor Lewis pointed out that this was a surprising result, since the ozone molecule is presumed to have no odd electrons.[3]

Becquerel's experiment, although only briefly described in his paper, was particularly simple and clear. A small hollow glass bar filled with air was suspended from its center horizontally by a gold wire, in a vessel between the poles of a magnet, the rest position of the bar making an angle of about 45° to the axis of the poles. In a magnetic field a couple will act upon the little bar, its magnitude depending on the strength and inhomogeneity of the field and on the difference between the susceptibility of the bar and of the surrounding medium. The angle through which the top of the wire had to be twisted to restore the bar to its original position in the presence of the field was measured in vacuum, in oxygen, and in two samples of ozonized oxygen. The torque in oxygen was much less than in vacuum, and in both ozone-oxygen mixtures it was still less, and was the least in the sample of higher concentration. Oxygen is paramagnetic and, therefore, from these data it appears that ozone is still more so. Quantitatively, assuming with Becquerel a value of ten volume per cent for his most concentrated sample, one finds the volume susceptibility of ozone to be more than three times that of oxygen.

So far as the author is aware no further work has been done on the magnetic properties of ozone. The experiment of Becquerel is described above in some detail, because its results differ radically from those obtained in the present experiment.

In the present work the Gouy method[4] in its simple form was used. A cylindrical armature 20 cm. long and 8 mm. diameter, depending from one side of an analytical balance, hung vertically so that its lower end rested in the homogeneous field between the pole faces of the large electromagnet of this laboratory. The upper end was then practically out of the field. The force exercised by the field on the armature was measured with the balance when the armature was surrounded by air, by oxygen, and by an ozone-oxygen mixture.

To secure these conditions the armature hung in a vertical glass cylinder

of 18 mm. diameter, open at the top through a short length of 5 mm. bore tube, through which the suspending fibre passed. Oxygen or the ozone-oxygen mixture entered at the bottom of the cylinder, passing upward and out freely at the top. Air from the room was drawn into the cylinder by aspirating slowly at a point in the glass tube line connecting the ozone generator with the glass cylinder in the magnet.

The ozone was prepared electrolytically, the concentrations used in this work being 6–7% by volume. Pure oxygen was obtained by simply decomposing the ozone in the ozone-oxygen mixture by the passage of the mixture over a spiral of resistance wire heated to redness. Thus the generator was a source of oxygen or of ozone, depending on whether the spiral was at red heat or at room temperature.

Two armatures were used, both of nearly the same dimensions, the first composed entirely of pyrex glass and the second being a shell of soda-lime glass filled with oxygen at atmospheric pressure. The former, which was used in the first experiments, was quite strongly diamagnetic, while the latter, used in the last experiments, proved to be slightly more paramagnetic than a similar volume of oxygen.

The first experiments showed a difference between oxygen and the ozone-oxygen mixture, this being, however, only very little more than the errors in measurement. The difference was such as to indicate that ozone was less paramagnetic than oxygen; that is, the force on the diamagnetic armature in the mixture was less than it was in pure oxygen. With stronger fields the difference increased, but small variations in the field produced fluctuations in the total force in either gas of the same order of magnitude as the difference, which was a few tenths of a milligram. There appeared, nevertheless, a tendency for the force to be less in the ozone mixture than in pure oxygen, and in the first three experiments at three different field intensities the difference was greater than the difference calculated for the gases assuming the susceptibility of ozone to be zero.

The volume susceptibility of oxygen was taken as expressed by the formula

$$\kappa = \frac{1.221 \times 10^{10}}{T^2}, \text{ and of air by } \kappa = \frac{0.256 \times 10^{10}}{T^2}$$

at 1 atmosphere pressure.[5] Two values of the susceptibility of the armature were had from two sets of measurements of the force in oxygen and in air, the first set at a low field strength yielding a less accurate result than the second. The values were -0.62×10^{-6} and -0.69×10^{-6}. The latter one has been used in these calculations, but either might be used without affecting the conclusions.

In view of the smallness of the effect and the magnitude of the errors of measurement, two experiments were performed on different days in which

a series of observations were made in each experiment on both oxygen and the ozone mixture, with the purpose of gaining an idea of the magnitude of the effect relative to the experimental errors. The results of these are given in figure 1. A smooth curve, No. 1 of the theoretical shape, that is, one in which the force varies as the square of the field intensity, has been drawn through the oxygen points. Relative to this, two calculated curves have been placed in the figure, No. 3 being that which Becquerel's results would indicate for this ozone concentration which was practically the same in the two experiments, and No. 2 that which should result if the suscepti-

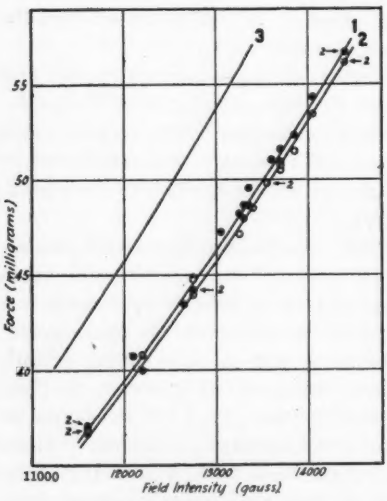


FIGURE 1

● Oxygen; ○ Ozone-Oxygen mixture.

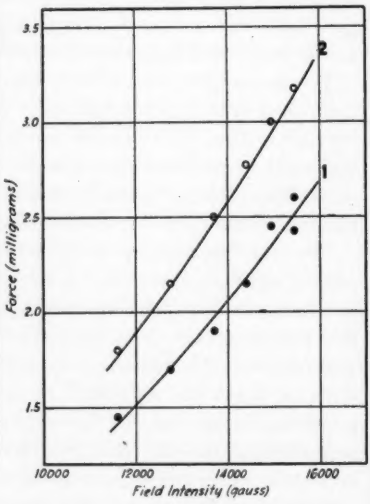


FIGURE 2

● Oxygen; ○ Ozone-Oxygen mixture.

bility of ozone were zero. In some cases observations fell at practically or identically the same point and are represented by a single point. These cases are indicated in the figure. Clearly the results are entirely different from those of Becquerel, and the points for ozone tend to lie below those for pure oxygen by an amount of the same order of magnitude as would be the case if the susceptibility of ozone were zero. It is also clear, however, that the results show fluctuations comparable with the effect itself. This is due chiefly to fluctuations in the field and to a lesser degree to the change in weight of the armature with changing room conditions and to direct errors of setting and reading.

To reduce the errors due to fluctuations in the field another armature was used, the second one described above, the attempt being made to have it of nearly the same susceptibility as oxygen. As mentioned, it was

slightly more paramagnetic than oxygen. The total force in oxygen was now small, and hence errors due to fluctuations in the field were correspondingly small. Three experiments on three different days, each consisting of a series of readings in oxygen and in an ozone-oxygen mixture, were performed with this armature, the technique of the measurements being successively improved. The results are shown in figures 2, 3 and 4. In each case a smooth curve, No. 1, of the theoretical shape, has been drawn through the oxygen points, and then a computed curve, No. 2, in which the susceptibility of ozone has been taken as zero, has been placed

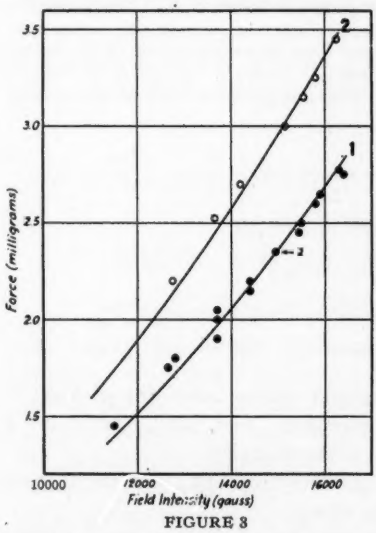


FIGURE 3

● Oxygen; ○ Ozone-Oxygen mixture.

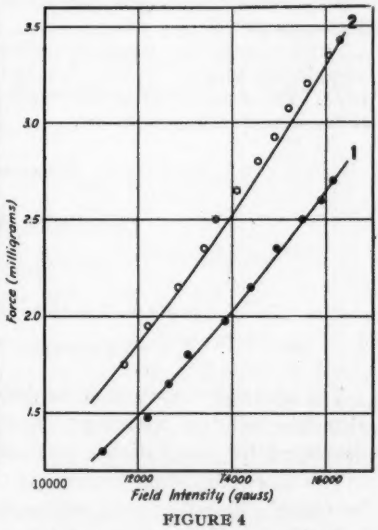


FIGURE 4

● Oxygen; ○ Ozone-Oxygen mixture.

upon the figure. In each case the curve required by Becquerel's results lies below the oxygen curve, and almost entirely off the figure.

Three sets of data for air and oxygen on these same days gave 0.174×10^{-6}, 0.174×10^{-6} and 0.173×10^{-6} for the volume susceptibility of the armature.

The results show that ozone is far from being more paramagnetic than oxygen. The tendency for the observations to lie above the computed curve in figures 2, 3 and 4 is evidence that ozone is diamagnetic, but the accuracy of these measurements does not permit a definite determination of the sign of the susceptibility. We can, however, conclude that the numerical value of the volume susceptibility of ozone is but a small fraction of that of oxygen. Thus the anomalous high paramagnetism of ozone recorded in the literature appears not to exist.

The author is glad to acknowledge his indebtedness to Professor G. N. Lewis for his interest and helpful suggestions; to Dr. Simon Freed for his kind advice throughout the work in a field with which the author was not familiar; and to Beatrice Wulf with whom much of the experimental work was done.

* NATIONAL RESEARCH FELLOW IN CHEMISTRY.

[1] Becquerel, *Compt. rend.*, **92**, 348 (1881).

[2] Schumeister, *Sitz. Akad. Wiss. Wien*, II, **83**, 45 (1881).

[3] Lewis, *Valence and the Structure of Atoms and Molecules*, Chemical Catalog Co., **1923**, pages 130, 147 et seq; *Chemical Reviews*, **1**, 234 (1924).

[4] See, for example, Stoner, *Magnetism and Atomic Structure*, E. P. Dutton and Co., **1926**, page 40.

[5] The expression for oxygen was calculated from an average value of the volume susceptibility using the three recent values listed by Stoner in Table IX of his book, see ref. 4. The expression for air was obtained from this and from the composition of air, taking account of the oxygen only.

GENERAL ARITHMETIC

BY MORGAN WARD

CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA

Communicated October 15, 1927

The abstract theory of a mathematical system consisting of a set of elements and two operations "multiplication," and, later, "addition" is developed by postulational methods with examples.

The more important results are the following. An "arithmetic" may be roughly described as a system in which

1. Every element is completely specified by a finite number of cardinal numbers.

2. "Division" is not always possible, and we can find when one element divides another element in a finite number of steps.

3. Unique resolution into "prime factors" is always possible.

For the case when multiplication is commutative we replace these vague requirements by an exact set of postulates, the necessary and sufficient conditions for a system to form an arithmetic. We give a general theory of recurring sequences, exhibiting the connections with the Dedekind field theory which we develop following Kronecker, as an arithmetic of forms without assuming the so-called "fundamental theorem of algebra." We next show that all systems of ideals and ideal numbers are abstractly equivalent and may be replaced by the system of rational integers, making ideals and ideal numbers unnecessary in algebraic arithmetic.

An arithmetic may be defined over any arbitrary class of distinct de-

numerable elements, and conversely any arithmetic determines such a class, giving connections with the algebra of logic. These connections are destroyed if we assume "multiplication" is not commutative.

For the case when multiplication is not a commutative, we replace "division" by "left division" and "right division," with analogous changes for other relations such as "equivalence." We then develop the theory of the greatest common divisor, least common multiple, equivalence with respect to unit factors and so on. It is shown that for an arithmetic, we must assume the units of the system are commutative with all the other elements of the system. Any "arithmetic" may be converted into a group by adjunction, and the theory of congruence and the fundamentals of Kronecker's theory of forms carry over almost unchanged. The complete development will be published shortly in a mathematical journal.

ON THE STRUCTURE OF A PLANE CONTINUOUS CURVE[1]

BY W. L. AYRES

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PENNSYLVANIA

Communicated October 11, 1927

If x and y are distinct points of a continuous curve M, the set of all points $[z]$, such that z lies on some arc of M with end-points x and y, is called the *arc-curve* xy and is denoted by $M(x + y)$. This was defined in a previous paper, "Concerning the Arc-Curves and Basic Sets of a Continuous Curve."[2] Among other results this paper contained the following theorems for the case where M lies in a plane:

A.—The arc-curve xy is itself a continuous curve.

B.—If K is a maximal connected subset of $M - M(x + y)$, then K has only one limit point in $M(x + y)$.

C.—If P is a point of a set K such that $K - P$ is the sum of two non-vacuous mutually separated sets K_1 and K_2, and A and B are distinct points of $K_1 + P$, then no point of K_2 is a point of any arc whose end-points are A and B and which lies wholly in K.

It is the purpose ot this paper to characterize types of points of a continuous curve and types of continuous curves by a set which is the limit of the arc-curve xy as y approaches x. Throughout the paper the letter M is used to denote a plane continuous curve and all point sets mentioned are considered as subsets of M. The theorems listed above will be referred to as "Theorem A," etc.

THEOREM 1.—*If the point x lies on no simple closed curve of M and α is any arc of M whose end-points are x and any other point z of M, then x is a limit point of the points of M which lie on α and separate x and z in M.*[3]

Proof. Suppose there exists an arc α of M whose end-points are x and z and such that x is not a limit point of the points of α which separate x and z in M. Then there exists a point y of α distinct from x and z such that no point of the subarc xy of α separates x and z in M. Let $[d]$ be the set of maximal connected subsets of $M-xy$, and for any maximal connected subset d let R_{xd} and R_{yd} be the first and last limit points of d on the subarc xy of α in the order x to y. If $R_{xd} \neq R_{yd}$, there exists an arc A_d, whose end-points are R_{xd} and R_{yd} and which lies in d except for these points.[4] Since the point x lies on no simple closed curve of M, $R_{xd} \neq x$ unless $R_{yd} = x$. If w is any point of $xy - x$, there exists a maximal connected subset d_w, such that w is between R_{xd_w} and R_{yd_w} on the arc xy if $w \neq y$, and if $w = y$, $R_{yd_w} = w \neq R_{xd_w}$.

Since only a finite number of the sets of $[d]$ have a diameter greater than any given positive number,[5] there exists a set d_1 of $[d]$, such that (1) $R_{yd_1} = y$ and (2) if d' is any set such that $R_{yd'} = y$, then $R_{xd'}$ lies on the subarc yR_{xd_1} of xy. In general, there exists a set d_n $(n > 1)$ of $[d]$, such that (1) $R_{xd_{n-1}}$ lies between R_{yd_n} and R_{xd_n} on xy and (2) if d' is any set of $[d]$ such that $R_{xd_{n-1}}$ lies between $R_{xd'}$ and $R_{yd'}$ on xy, then $R_{xd'}$ lies on the subarc yR_{xd_n} of xy.

We have the order $y, R_{yd_2}, R_{xd_1}, R_{yd_3}, R_{xd_3}, R_{yd_4}, R_{xd_2}, \ldots, x$ on the arc xy. All of the points of this series are distinct except that for any i we may have $R_{xd_i} = R_{yd_{i+2}}$. Using the fact that each time d_n is chosen to satisfy condition (2), it is not difficult to see that x is the sequential limit point of $R_{xd_1}, R_{xd_2}, \ldots$. If $R_{xd_i} \neq R_{yd_{i+2}}$, let $R_{xd_i}R_{yd_{i+2}}$ denote the subarc of xy which has these two points as end-points, and if $R_{xd_i} = R_{yd_{i+2}}$, let $R_{xd_i}R_{yd_{i+2}}$ denote this single point. Let yR_{yd_2} denote the subarc of xy whose end-points are y and R_{yd_2}. Let

$$\alpha_1 = x + A_{d_1} + R_{xd_1}R_{yd_3} + A_{d_3} + \ldots + A_{d_{2i-1}} + R_{xd_{2i-1}}R_{yd_{2i+1}} + \ldots,$$
$$\alpha_2 = x + yR_{yd_2} + A_{d_2} + R_{xd_2}R_{yd_4} + A_{d_4} + \ldots + A_{d_{2i}} + R_{xd_{2i}}R_{yd_{2i+2}} + \ldots.$$

Since α_j $(j = 1, 2)$ is connected, x is the sequential limit point of $[R_{xd_i}]$ and the diameter of A_{d_i} approaches zero as i increases indefinitely,[5] the set α_j is an arc of M whose end-points are x and y. The arcs α_1 and α_2 have only x and y in common; therefore, $\alpha_1 + \alpha_2$ is a simple closed curve of M containing x, which is contrary to hypothesis.

Since every point of M is either a cut-point or a non-cut point, and non-cut points which lie on no simple closed curve are end-points,[6] we have the following.

COROLLARY. *If P is a cut-point of M which lies on no simple closed curve of M or if P is an end-point of M, then P is a limit point of the cut-points of M.*

THEOREM 2.—*If x is a point of M and C_x denotes the set consisting of x and all points $[y]$ such that M contains a simple closed curve containing both x and y, then C_x is a continuous curve.*

Proof. Obviously C_x is connected, and by an argument similar to that used in proving theorem 1 we may show that C_x is closed. It remains to see that C_x is connected *im kleinen*. If P is a point of C_x and ϵ is a positive number, there exists a number $\delta_\epsilon > 0$ such that every point of M within a distance δ_ϵ of P can be joined to P by an arc of M, every point of which is within a distance ϵ of P. Let Q be any point of C_x within a distance δ_ϵ of P and let PQ denote an arc of M with end-points P and Q and every point of which is within a distance ϵ of P. We shall show that every point of PQ belongs to C_x, which proves the theorem.

Suppose PQ contains a point Z which does not belong to C_x. Then $P \neq Z \neq Q$. Since C_x is closed, there are first points P_1 and P_2 of C_x on the subarcs ZP and ZQ of PQ in the order Z to P and Z to Q. If either P_1 or P_2, say P_1, is the point x, the arc P_1ZP_2 of PQ plus either of the arcs from x to P_2 of a simple closed curve of M containing both x and P_2, is a simple closed curve of M containing both x and Z, and thus Z belongs to C_x. If neither P_1 nor P_2 is x, let J_1 and J_2 denote simple closed curves of M which contain P_1 and x and P_2 and x, respectively. If J_1 contains P_2, then the arc of J_1 from P_1 to P_2 which contains x together with the arc P_1ZP_2 of PQ is a simple closed curve of M containing x and Z. We proceed similarly if J_2 contains P_1. If J_1 does not contain P_2 and J_2 does not contain P_1, let α denote one of the arcs of J_2 from P_2 to x and let P_3 be the first point of J_1 on α in the order P_2 to x. If $P_3 = x$, then α plus the arc P_1ZP_2 of PQ plus either of the arcs of J_1 from P_1 to x is a simple closed curve of M containing x and z. If $P_3 \neq x$, then the subarc P_3P_2 of α plus the arc P_1ZP_2 of PQ plus the arc of J_1 from P_1 to P_3 which contains x is a simple closed curve of M containing x and Z. Thus in any case the point Z belongs to the set C_x

THEOREM 3.—*If K is a maximal connected subset of $M - C_x$, K has only one limit point in C_x.*

With the use of a theorem due to Wilder,[4] the proof of theorem 3 follows that of theorem 2.

THEOREM 4.—*If K is a maximal connected subset of $C_x - x$, then $K + x$ is a cyclicly connected[7] continuous curve.*

Proof. From theorem 2 we see that $K + x$ is a continuous curve. Suppose $K + x$ contains a cut-point P, that is, $K + x - P$ is the sum of two non-vacuous mutually separated sets K_1 and K_2. Since K is connected, P is not the point x. Then x lies in one of the two sets K_1 or K_2 and let y be a point of the other. Since y belongs to C_x there exist two arcs of M from y to x which have only their end-points in common. But this is impossible, since every arc from y to x must contain P. Then $K + x$

contains no cut-point and by a theorem due to G. T. Whyburn,[8] $K + x$ is a cyclicly connected continuous curve.

THEOREM 5.—*If the continuous curve M contains a cut-point and a simple closed curve, it contains a cut-point which lies on a simple closed curve of M.*

Theorem 5 may be established by the methods of theorem 1.

THEOREM 6.—*If u and v are distinct points of $M(x + y)$, then $M(u + v)$ is a subset of $M(x + y)$.*

This theorem follows easily from theorem B.

Definition. Let x be a point of M and $y_1, y_2, y_3, \ldots$ be a sequence of points of M distinct from x, such that the distance from x to y_i approaches zero as i increases indefinitely. Let K denote the set of all points which are common to all the sets $M(x + y_i)$. The set K exists and is non-vacuous. If the set K is independent of the choice of the sequence of points $y_1, y_2, y_3, \ldots$ approaching x as a limit, then K will be called the *limiting arc-curve of M at the point x* and will be denoted by Lim $M(x)$.

THEOREM 7.—*If x is a point of M which lies on no simple closed curve of M, the limiting arc-curve at x exists and is the point x.*

Proof. If the theorem is not true, there exists a sequence of points $y_1, y_2, y_3, \ldots$ such that the sets $M(x + y_i)$ have in common a point $z \neq x$. By theorem 1, if α is an arc of M whose end-points are x and z then α contains a point p distinct from x and z such that p separates x and z in M. Then $M - p = M_x + M_z$, where M_x and M_z are non-vacuous mutually separated sets containing x and z, respectively. As x is not a limit point of M_z, there exists a number $\epsilon > 0$ so that no point of M_z is within a distance ϵ of x. There exists an integer n_ϵ such that if $i > n_\epsilon$, the point y_i lies within a distance ϵ of x. Then $M_x + p$ contains y_i for $i > n_\epsilon$, and by theorem C, $M(x + y_i)$ cannot contain z. But we assumed that $M(x + y_i)$ contains z for every i. Hence, for every sequence $[y_i]$ approaching x, the point x is the only point common to all the sets $M(x + y_i)$.

THEOREM 8.—*If x is a cut-point of M and lies on some simple closed curve of M, the limiting arc-curve at x does not exist.*

Proof. Suppose $M - x$ is the sum of two non-vacuous mutually separated sets M_1 and M_2 and that the limiting arc-curve at x exists. From theorem C it follows that if all points of the sequence $y_1, y_2, y_3, \ldots$ belong to M_1, the limiting arc-curve at x is a subset of $M_1 + x$. Similarly, if all points of the sequence belong to M_2, Lim $M(x)$ is a subset of $M_2 + x$. But the Lim $M(x)$ is independent of the choice of the sequence and thus it must be common to the sets $M_1 + x$ and $M_2 + x$; that is, if the limiting arc-curve exists at a cut-point x, it must consist of the single point x. But if x lies on some simple closed curve J of M and $[y_i]$ is a sequence, every point of which belongs to J, the arc-curve $M(x + y_i)$ contains J for every i. Thus J is a subset of the set of common points for this se-

quence. Then the set of common points depends on the choice of the sequence. Hence Lim $M(x)$ does not exist.

THEOREM 9.—*If x is a non-cut-point of M which lies on a simple closed curve of M, the limiting arc-curve at x exists and is the set C_x.*

Proof. As x is not a cut-point of M, $C_x - x$ is connected; then from theorem 4, C_x is a cyclicly connected continuous curve. It is not difficult to see that if y and z are any points of C_x distinct from x and from each other, there exists an arc of C_x which has end-points x and y and contains z. With this and theorem 3, we see that $M(x + y) = C_x$ for every point $y \neq x$ of C_x. Also if y is a point of a maximal connected subset D of $M - C_x$, then $M(x + y)$ is contained in $D + C_x$ and contains C_x. Then C_x is a subset of $M(x + y)$, if y is any point of M distinct from x. Now suppose there exists a sequence of points $[y_i]$ of M approaching x such that the arc-curves $M(x + y_i)$ have in common a point w which is not a point of C_x. Let D_w denote the maximal connected subset of $M - C_x$ containing the point w. By theorem 3, D_w has only one limit point in C_x, and since x is not a cut-point of M, this limit point is not x. There exists a number $\epsilon > 0$ such that no point of D_w is within a distance ϵ of x. There exists a number $n_\epsilon > 0$ such that for $i > n_\epsilon$, the distance from x to y_i is less than ϵ. For $i > n_\epsilon$, y_i must lie in $M - D_w$, and by theorem C, for $i > n_\epsilon$, $M(x + y_i)$ is a subset of $M - D_w$. Then the point w is not common to all the sets $M(x + y_i)$. Therefore, $C_x = \text{Lim } M(x)$.

Since a definite point of a continuous curve M must satisfy one, and not more than one, of the hypotheses of theorems 7, 8 and 9, we have

THEOREM 10.—*If x is a point of a continuous curve M, then (a) x lies on no simple closed curve of M, (b) x is a non-cut-point of M and lies on a simple closed curve of M, (c) x is a cut-point of M and lies on a simple closed curve of M, if and only if (a) Lim $M(x) = x$, (b) Lim $M(x) = C_x \neq x$, (c) Lim $M(x)$ does not exist.*

THEOREM 11.—*In order that a continuous curve M be acyclic,[9] it is necessary and sufficient that Lim $M(x) = x$ for every point x of M.*

Theorem 11 is a consequence of theorem 10(*a*).

THEOREM 12.—*In order that a continuous curve M be cyclicly connected, it is necessary and sufficient that M contain a point x such that Lim $M(x) = M$.*

Proof. That the condition is necessary follows from the property of cyclicly connected continuous curves stated above, namely, that if x and y are distinct points of such a curve, $M(x + y) = M$. We shall show that it is sufficient. If M is not cyclicly connected, M contains a cut-point P.[8] Then $M - P$ is the sum of two non-vacuous mutually separated sets M_1 and M_2, and we suppose that $M_1 + P$ contains the point x. Let y be any point of M_1 distinct from x. By theorem C, $M(x + y)$ is a subset of $M_1 + P$ and thus contains no point of M_2. Since $M(x + y)$ contains Lim $M(x)$, Lim $M(x) \neq M$.

It is evident that if M is cyclicly connected, then Lim $M(x) = M$ for *every* point x of M and conversely.

THEOREM 13.—*If M is a continuous curve, either* (1) *Lim $M(x) = x$ for every poin. x of M, or* (2) *Lim $M(x) = M$ for every point x of M, or* (3) *M contains a point x such that the limiting arc-curve of M at x does not exist.*

Proof. If (1) does not hold, then M contains a simple closed curve by theorem 11. If (2) is not true, M is not cyclicly connected by theorem 12. Hence M contains a cut-point.[8] By theorem 5, M contains a cut-point which lies on a simple closed curve of M. By theorem 8, the limiting arc-curve of M does not exist at this point. Thus, if neither (1) nor (2) is true, (3) must be true.

[1] Presented to the American Mathematical Society, May 7, 1927.

[2] This paper has been submitted for publication in *Trans. Amer. Math. Soc.*

[3] A point P of a set H is said to separate two points x and y in H if H-P is the sum of two mutually separated sets, one containing x and the other containing y.

[4] Wilder, R. L., *Fund. Math.*, **7**, 1925 (342).

[5] Cf. my paper, "Note on a Theorem Concerning Continuous Curves," *Annals Math.*, **28**, 1927 (501–2).

[6] Ayres, W. L., *Annals Math.*, **28**, 1927 (396–418).

[7] A continuous curve H is said to be cyclicly connected if and only if every two points of H lie together on some simple closed curve which is a subset of H. Cf. Whyburn, G. T., these PROCEEDINGS, **13**, 1927 (31–38).

[8] Loc. cit., theorem 1.

[9] A continuous curve is said to be acyclic if it contains no simple closed curve. This term has been introduced recently by H. M. Gehman.

ON THE ARITHMETIC OF ABELINA FUNCTIONS

BY E. T. BELL

CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated October 15, 1927

1. Systematic application of abelian functions to the theory of numbers discloses many new arithmetical phenomena of considerable interest. Particularly is this the case when the periods of the functions are connected by one or more singular relations in the usual sense. An example is afforded by the novel type of arithmetical invariance made precise in §2, and illustrated in §3, which, roughly, is as follows.

Consider a set of $n > 1$ quadratic forms in the same s indeterminates. The sets of rational integral values of the s indeterminates which represent an arbitrary set of n integers representable simultaneously in the n forms are separated into sets of residue classes with respect to a modulus α. If a particular set of h of these classes is such that the number of repre-

sentations of the specified set of n integers which fall into each of the h classes is finite, and if further the respective numbers of representations in the h classes are $w, k_2w, \ldots, k_2w$, where the k's are absolute constants, the set of positive integers $1, k_2, \ldots, k_3$ may legitimately be considered as an arithmetical invariant of the set of n forms with respect to α. Next, all of the sets of n arbitrary integers simultaneously representable in the n forms are separated into sets of residue classes with respect to a modulus μ, and in general μ is taken different from α. Each of certain of such sets yields, as above, an invariant set of integers with respect to α; a set of any number of such invariant sets is an invariant set of the n forms with respect to α and μ. Obviously the number of invariant sets is finite, if any exist; not every set of n forms generates an invariant set, nor does the existence of a set for a particular pair of values of α, μ imply the existence of a set for another pair of values.

The remarkable situation just described typifies a cardinal distinction between the arithmetic implicit in the abelian functions and that in the elliptic: simultaneous representations and their properties are the rule in the abelian case, while in the elliptic they are the exception.

The examples in §3 are among the simplest illustrations of §2. The first was deduced from the rational algebraic relations connecting singular functions of two arguments, for both arguments zero, when the singular relation between the periods has the invariant 5; the second is equivalent to one of the classic biquadratic relations.

2. A one-rowed matrix $(n_1, \ldots, n_p)$ of p rational integers will be called integral of order p. If δ is a constant integer different from zero, and r_j is the least positive residue modulo δ of n_j $(j = 1, \ldots, p)$, we write

$$(n_1, \ldots, n_p) \equiv (r_1, \ldots, r_p) \bmod \delta,$$

and call $(r_1, \ldots, r_p)$ the residue mod δ of $(n_1, \ldots, n_p)$. Such residues being matrices, their equality is defined. The residue class $C(r_1, \ldots, r_p)$ mod δ is the set of all integral matrices of order p each of which has the residue $(r_1, \ldots, r_p)$ mod δ; the number of these classes is δ^p. It is convenient presently to use both ( ) and $\{\ \}$ to designate integral matrices.

Henceforth α, μ are constant integers different from zero, and the n distinct quadratic forms Q_j,

$$Q_j \equiv Q_j(u_1, \ldots, u_s) \qquad (j = 1, \ldots, n),$$

in the s indeterminates $u_1, \ldots, u_s$, with rational integral coefficients, are given and fixed.

If for $(m_1, \ldots, m_n)$ the general element (integral matrix of order n) in $C(\mu_1, \ldots, \mu_n)$ mod μ there exist s rational integers $a_1, \ldots, a_s$ such that

$$Q_j(a_1, \ldots, a_s) = m_j \quad (j = 1, \ldots, n),$$

we say that $(m_1, \ldots, m_n)$ is equivalent to $\{a_1, \ldots, a_s\}$ with reference to $[Q_1, \ldots, Q_n]$, and write

$$(m_1, \ldots, m_n) \sim \{a_1, \ldots, a_s\} \text{ ref. } [Q_1, \ldots, Q_n] \qquad (1)$$

Keep $(m_1, \ldots, m_n)$ fixed, and therefore also $(\mu_1, \ldots, \mu_n)$, its residue mod μ, and consider the set of all $\{a_1, \ldots, a_s\}$ satisfying (1). If in this set there are only a finite number of matrics, necessarily all unequal, they will fall into $h \leqq \alpha^s$ distinct residue classes K_i mod α, such that K_i contains precisely w_i values of $\{a_1, \ldots, a_s\}$ satisfying (1), where w_i is finite $(i = 1, \ldots, h)$. If now $h - 1$ of the integers w_i $(i = 1, \ldots, h)$ are constant integral multiples of the remaining one, which, without loss of generality, may be taken to be w_i, so that

$$w_i = k_i w_1 \ (i = 1, \ldots, h), \quad k_1 = 1,$$

where the k's are absolute constants, we shall say that $M \equiv (\mu_1, \ldots, \mu_n)$, the residue of the general element $(m_1, \ldots, m_n)$ of $C(\mu_1, \ldots, \mu_n)$ mod μ, is finite in the class matrix $K \equiv ((K_1, \ldots, K_h))$ mod α of order h, and has with respect to K the relative weight $W \equiv \mid 1, k_2, \ldots, k_h \mid$. We shall symbolize this (exceptional) situation thus,

$$M \text{ resp. } K = W. \qquad (2)$$

Next, in (1), let $(m_1, \ldots, m_n)$ range over the set of general elements in the respective μ^n residue classes mod μ of integral matrices of order n. Now it may happen, and in general does for the $[Q_1, \ldots, Q_n]$ arising from singular abelian functions, that several of the residues mod μ are finite in certain class matrices mod α, and that for each such residue mod μ there is a weight relation of the type defined by (2). Let

$$M^{(j)} \text{ resp. } K^{(j)} = W^{(j)} \ (j = 1, \ldots t) \qquad (3)$$

be any set of these relations for the $M^{(j)}$ all different. Then, subsuming the t relations (3) under a single matric relation, we write

$$(M^{(1)}, \ldots, M^{(t)}) \text{ resp. } ((K^{(1)}, \ldots, K^{(t)})) = \mid W^{(1)}, \ldots, W^{(t)} \mid \qquad (4)$$

which may be read: the residue matrix $(M^{(1)}, \ldots, M^{(t)})$ mod μ is finite in the class matrix $((K^{(1)}, \ldots, K^{(t)}))$, mod α and has with respect to it the relative weight $\mid W^{(1)}, \ldots, W^{(t)} \mid$. When $t = 1$, (4) becomes (2). In (3) the orders of $K^{(j)}$ $(j = 1, \ldots, t)$ are not necessarily equal, although in the examples of §3 they are.

From the above definitions it is clear that a relative weight is a set of absolute constants (positive integers) connected with the simultaneous representation of sets of n arbitrary integers in the given set of n quadratic forms in s indeterminates; in this sense a relative weight is an invariant of $[Q_1, \ldots, Q_n]$ determined by α, μ alone.

3. Take $s = 6$, $n = 2$, $\alpha = 2$, $\mu = 8$,

$$Q_1 \equiv u_1^2 + u_2^2 + u_3^2 + u_4^2 + u_5^2 + u_6^2,$$

$$Q_2 \equiv 2(u_1u_6 + u_2u_5 + u_3u_4) + u_4^2 + u_5^2 + u_6^2,$$

and consider (1), (2) of §2 in this case for the following 6 values M_j of the possible 64 of $M \equiv (\mu_1, \mu_2)$,

$$\begin{array}{lll} M_1 = (3,3), & M_3 = (3,0), & M_5 = (6,1), \\ M_2 = (7,7), & M_4 = (7,0), & M_6 = (6,5), \end{array}$$

with respect to the 6 residue classes, I, II..., VI, of the 64 possible $C\{u_1, \ldots, u_6\}$ mod 2,

$$\begin{array}{r} \text{I} : (0, 1, 1, 1, 0, 0), \\ \text{II} : (0, 0, 0, 1, 1, 1), \\ \text{III} : (0, 0, 1, 1, 1, 0), \\ \text{IV} : (1, 1, 1, 0, 0, 0), \\ \text{V} : (0, 0, 1, 0, 1, 0), \\ \text{VI} : (1, 1, 1, 1, 1, 1). \end{array}$$

Corresponding to (3), we have the following table, in which (i, j, k) is either (1,3,5) or (2,4,6).

(μ_1, μ_2)	:	Class matrix	:	Relative weight
M_i	:	((I, II))	:	$\left\vert \begin{matrix} 1, 2 \end{matrix} \right\vert$
M_j	:	((III, IV))	:	$\left\vert \begin{matrix} 1, 2 \end{matrix} \right\vert$
M_k	:	((V, VI))	:	$\left\vert \begin{matrix} 1, 2 \end{matrix} \right\vert$,

which can be combined into a single relation of the type (4). For example, the third row of the table states that if m_1, m_2 are any rational integers such that either $m_1 \equiv 6$, $m_2 \equiv 1$ mod 8, or $m_1 \equiv 6$, $m_2 \equiv 5$ mod 8, then the number of sets of integers $\{u_1, \ldots, u_6\}$ satisfying $Q_1 = m_1$, $Q_2 = m_2$, in which all values of $u_1, \ldots, u_6$ are odd (see VI) is twice that in which only the values of u_3, u_5 are odd (see V).

As a second example, let c be an arbitrary constant integer which is not a square. Take $n = 2$, $s = 8$, $\alpha = 2$, $\mu = 4$. Of the 256 residue classes $C\{u_1, \ldots, u_8\}$ mod 2
the pair

$$\begin{array}{r} \text{I} : (1, 1, 1, 1, 0, 0, 0, 0), \\ \text{II} : (0, 0, 0, 0, 1, 1, 1, 1), \end{array}$$

gives for either of $M \equiv (\mu_1, \mu_2) = (0, 0), (0, 2)$, and the pair of forms

$$Q_1 \equiv u_1^2 + u_2^2 + u_3^2 + u_4^2 + c(u_5^2 + u_6^2 + u_7^2 + u_8^2),$$
$$Q_2 \equiv u_1u_8 + u_2u_7 + u_3u_6 + u_4u_5,$$

the following instance of (2),

$$M \text{ resp. } ((\text{I, II})) = \left| \begin{matrix} 1, 1 \end{matrix} \right|,$$

That is, the number of sets $\{u_1, \ldots, u_8\}$ satisfying $Q_1 = m_1$, $Q_2 = m_2$. in which $u_1, \ldots, u_8$ have the respective residues mod 2 indicated in I, is equal to the number of sets in which the residues are as in II, for either $m_1 \equiv m_2 \equiv 0 \bmod 4$ or $m_1 \equiv 0$, $m_2 \equiv 2 \bmod 4$.

Both examples illustrate the following for a pair ($n = 2$) of quadratic forms in $2r(= s)$ indeterminates. Let p,q be rational integers such that $p^2 + 4q$ is not a square. Then, $\sum$ referring to $j = 1, \ldots, r$, a typical pair of forms for which relations of the type (4) subsist, when α, μ are powers of 2, is

$$Q_1(u_1, \ldots, u_{2r}) = \sum(u_j^2 + qu_{2r-j+1}^2),$$
$$Q_2(u_1, \ldots, u_{2r}) = \sum u_{2r-j+1} \quad (2u_j + pu_{2r-j+1}).$$

In the first example $r = 3$, $p = q = 1$; in the second, $r = 4$, $p = 0$, $q = c$. since evidently $Q_2 = m_2$ is equivalent to $2Q_2 = 2m_2$.

GROUPS WHOSE OPERATORS ARE OF THE FORM $s^p t^q$

By G. A. Miller

Department of Mathematics, University of Illinois

Communicated October 17, 1927

A. Cayley called attention to the fact that if a group G is generated by two non-commutative operators s and t which satisfy the equation $st = t^2s^2$ then it is not possible to represent all of the operators of G by the expression $s^p t^q$.[1] This tends to augment the interest in the question what properties a group which is generated by the two operators s and t must have in order that all the operators of this group are of the form $s^p t^q$, where p and q are integers which do not exceed the orders of s and t, respectively. In what follows the symbol G will represent such a group and the subgroups generated by s and t will be represented by S and T, respectively. If a subgroup of S is invariant under t it is obviously invariant under G, and if a subgroup of T is invariant under s it is also invariant under G. We shall, therefore, represent by S_1 and T_1 the maximal subgroups of S and T which are invariant under t and s, respectively. The index of S_1 under S will be represented by m and the index of T_1 under T will be represented by n.

Since all the operators of G can be represented by $s^p t^q$ they can also be represented by $t^q s^p$, for the inverses of all the operators of a finite group

represent each operator of this group once and only once. Hence, we may assume without loss of generality that $m \leqq n$. When $m = 1$ it is clear that all the operators of G can be written in the form $s^p t^q$ and hence we shall assume for the present that $m > 1$. It will soon appear that this assumption gives rise to a contradiction. It is clear that the order of G is a divisor of the product of the orders of s and t. The quotient group of G with respect to its invariant subgroup generated by S_1 and T_1 is generated by two operators s_1, t_1 of orders m and n, respectively, corresponding to the operators s and t, respectively. Since all the operators of G are of the form $s^p t^q$ all the operators of this quotient group are of the form $s_1^{p_1} t_1^{q_1}$. The operators of this quotient group can, therefore, be written in the form of a rectangle of which the first row is composed of the operators generated by t_1. Since T_1 is a maximal invariant subgroup of T under G it results that the group generated by t_1 involves no invariant subgroup under s_1 besides the identity. Hence a conjugate of t_1 under s_1 generates a group which has only the identity in common with the group generated by t_1. This is contrary to the well-known theorem that the index of the cross-cut of any two distinct conjugate subgroups under one of these subgroups is always less than the index of these subgroups under the entire group.[2]

As we have arrived at a contradiction by assuming that $m > 1$ it results that *a necessary and sufficient condition that all the operators of a group which is generated by s and t can be represented in the form $s^p t^q$ is that at least one of these operators is transformed into a power of itself by the ether.* If only one of them is thus transformed its order must exceed the index of the largest invariant subgroup involved in the group generated by the other. It is now easy to see why the condition considered by A. Cayley is impossible, for if the conditions $st = t^2 s^2$ is satisfied neither of the two operators s and t can transform the other into a power of itself unless each is the inverse of the other. That is, a necessary and sufficient condition that all the operators of a group generated by s and t can be written in the form $s^p t^q$ when the operators satisfy the equation $st = t^2 s^2$ is that each of these operators is the inverse of the other. This fact was noted by A. Cayley but he did not find a necessary and sufficient condition that all the operators of a group generated by s and t can be represented in the form $s^p t^q$.

[1] A. Cayley, *Messenger of Mathematics*, vol. **7** (1878), p. 188.

[2] Miller, Blichfeldt, Dickson, *Theory and Applications of Finite Groups*, **1916**, p. 68.

THE CONTINUOUS SPECTRUM OF HYDROGEN

By Joseph Kaplan[1]

Palmer Physical Laboratory, Princeton University

Communicated October 13, 1927

There are two known continuous spectra of hydrogen, one on the limit of the Balmer series and explained by Debye according to the Bohr theory and one extending from the Schumann region far into the visible. The latter has received various explanations the best and more recent of which seem to be the ones advanced by Blackett and Franck[2] and extended by Oldenberg.[3] According to this explanation the continuous spectrum is due to the dissociation of excited hydrogen molecules, the initial states being ones in which the molecule possesses electronic energy with or without vibration and rotation and the final states ones in which the molecule has been dissociated into two neutral atoms with various kinetic energies relative to each other. The long wave-length limit of a band series corresponds, therefore, to the short wave-length limit of a continuous spectrum and arises in transitions in which the resulting product is two neutral atoms having no relative kinetic energy. The continuous part of the spectrum is emitted in cases where the two atoms do possess kinetic energy.

It is seen from the above that the energy of the molecule in the excited state must be enough to provide for the energy of dissociation, the kinetic energy of the atoms relative to each other and for the energy which is radiated as continuous spectrum. That a similar process occurs in absorption has been shown by Dieke and Hopfield.[4] The work of Witmer[5] on the analysis of bands in the Lyman region gives as the heat of dissociation a value of 4.34 volts and that found by Dieke and Hopfield was 4.38 volts.

The purpose of this note is to call attention once more and to explain a continuous spectrum in hydrogen that has been observed by Horton and Davies[6] and by Crew and Hulburt[7] and others, and possessing the peculiar characteristic of occurring without the presence of either atomic or molecular lines. Horton and Davies observed that hydrogen at pressures less than 1 mm. of mercury became suffused with a blue glow when excited by electrons of less than 15 volts' velocity. This glow was a continuous spectrum showing no trace of atomic or molecular lines.

Hulburt and Crew observed what seemed to be the same spectrum in the first striation of a discharge tube at 1 mm. of mercury. Bramley[8] has found that at 1 mm. pressure the electrons in the striation nearest the cathode have velocities of about 10 or 15 volts and Horton and Davies[9] give 12.6 volts as the lowest value of the electron velocities necessary to give this spectrum and find that a slight increase in voltage causes the blue

glow to change to a pink one which shows atomic and molecular lines. All of these facts are to be kept in mind since they unite in giving a simple explanation as to the origin of the spectrum which Crew and Hulburt[10] regard as a retardation spectrum of H_3^+ similar to the spectrum of H^+ on the short wave-length side of the Balmer series. They cite as one of their arguments in favor of this explanation the strong concentrations of H_3^+ found by positive ray analysis in this region.

Now the longest "short wave-length" limit of any of the continuous spectra of hydrogen will correspond to a transition in which the molecule is dissociated into two neutral atoms possessing no relative kinetic energy. From their measurements on the absorption spectrum of H_2, Dieke and Hopfield[11] give 11.1 volts as the excitation potential of this level. Subtracting the 4.4 volts necessary to dissociate the molecule one is left with 6.7 volts which is given off as radiation. This gives 1830 A.U. as the longest short wave-length limit. This transition probably does not occur because Oldenberg[12] observed no continuous spectrum when he excited molecular hydrogen in the presence of argon. The rôle of the argon was to limit the excitation to 11.6 volts. The reason given by Oldenberg for his failure to observe a continuous spectrum in this case was that the amount of vibrational energy necessary for dissociation was too large for the process to occur. For this reason, therefore, one should expect this continuous spectrum to have a short wave-length limit at about 1700 A.U. since the amount of energy radiated will be certainly no less than $11.6-4.4 = 7.2$ volts.

Lewis[13] photographed the continuous spectrum of hydrogen with a fluorite spectrograph and found a short wave-length limit at about 1750 A.U. and also observed the reappearance of the many line spectrum at about 1700. This is in sufficiently good agreement with the above predicted value to warrant the suggestion that the blue spectrum arises when the molecule, in the first electronic state and possessing more than $11.6-11.1 = 0.5$ volt of vibrational energy, splits into two atoms. Hopfield[19] found 1680 A.U. as this limit. Using this value we obtain 0.58 volt as the necessary initial vibrational energy of the molecule for the above process to occur.

The fact that no lines are to be observed is expected since none of the radiation arising in transitions from this state to lower ones lies in the visible or in the part of the ultra-violet accessible to the quartz spectrograph. Theoretically the smallest quantum jump possible is that from the 11.1 electronic level to that vibration level in which the vibrational energy of the molecule is just short of that necessary for dissociation. Actually the smallest jump is somewhat larger than this in view of the experiments of Oldenberg with argon and hydrogen which was mentioned above. In view of this experiment the smallest jump is certainly equal to

or greater than $11.6-4.4 = 7.2$ volts and this does not allow any lines above 1720 A.U., once more in good agreement with the theory.

An additional argument which may be cited in support of this theory is the fact that in the first striation in a hydrogen tube (which is usually dark blue and quite unique, the others being red) the electron energies vary from 10 to 15 volts agreeing with the value of about 11.6 volts necessary for the excitation of the continuous spectrum. Horton and Davies observed that if the electron velocities were increased above a certain critical value around 12.6 volts, the blue continuous spectrum changed to a pink one which showed also the Balmer lines and the secondary lines. The second electron level in H_2 is at 12.2 volts which is certainly too near the first one to give visible radiation. The next level, for which Witmer gives a radiation potential of 14.26 volts, is sufficiently high to account both for the secondary lines and the Balmer lines, which are no doubt emitted due to the excitation of the products of dissociation by electrons having nearly the same velocity as those which excite the molecule to the third electron level.

The fact that the first striation is blue has been long known and the interpretation given here for the spectrum of this striation is in agreement with the few investigations concerning first striations. Among others, Stark[14] arrived at the conclusion that the order of the maxima reached in the spectrum of the first striation is the same as the order of the excitation potentials of these spectra. When mercury is present, for example, the first striation is made up of green, blue and pink-white partial striations. The green one corresponds to the excitation of the mercury (which is the lowest), the blue mainly to that of this continuous spectrum of H_2 and the pink one to the excitation of Balmer and visible molecular lines. Since, according to the ideas presented here, the blue continuous spectrum is to be the one necessitating the lowest excitation potential, the fact that it appears in its proper place in the first striation is another argument in support of the theory.

Houtermans[15] gives a similar explanation for the continuous spectrum observed in the fluorescence of Hg vapour and refers to Pringsheim,[16] Mecke[17] and Kuhn[18] who found that the change in the binding when an electron jump occurs in the case of halogens can be so radical as to give a maximum 1400 Å away from the convergence limit of the bands. These observations support the possibility of the existence of a continuous spectrum which stretches as far as the one explained here, which necessitates relative kinetic energy of the atoms amounting to as much as four volts.

[1] NATIONAL RESEARCH FELLOW IN PHYSICS.

[2] Blackett and Franck, *Zeits. Physik*, **34**, 389 (1925).

[3] Oldenberg, *Ibid.*, **41**, 1 (1927).

[4] Dieke and Hopfield, *Zeits. Physik*, **40**, 299 (1926).
[5] Witmer, *Proc. Nat. Acad. Sci.*, **12**, 238 (1926).
[6] Horton and Davies, *Phil. Mag.* **46**, 872 (1923).
[7] Crew and Hulburt, *Phys. Rev.*, **28**, 936 (1926).
[8] Bramley, *Ibid.*, **26**, 794 (1925).
[9] Horton and Davies, loc. cit.
[10] Crew and Hulburt, loc. cit.
[11] Dieke and Hopfield, loc. cit.
[12] Oldenberg, loc. cit.
[13] Lewis, *Phys. Rev.*, **16**, 367 (1920).
[14] Stark, *Ann. Phys.*, **52**, 221, 225 (1917).
[15] Houtermans, *Zs. Phys.*, **41**, 140 (1927).
[16] Pringsheim, *Ibid.*, **5**, 130 (1921); and **7**, 216 (1921).
[17] Mecke, *Ann. Phys.*, **71**, 104 (1923).
[18] Kuhn, *Ibid.*, **39**, 77 (1926).
[19] Hopfield, *Phys. Rev.*, **20**, 573 (1922).

THE IMPULSE MOMENT OF THE LIGHT QUANTUM

BY ARTHUR EDWARD RUARK AND HAROLD C. UREY

YALE UNIVERSITY AND JOHNS HOPKINS UNIVERSITY

Communicated September 19, 1927

§1. For some years it has been customary to treat the light quantum as a particle and to emphasize those characteristics of the electron which point to its wave-like character. This idea, beautifully substantiated by the success of wave dynamics, has led Beck[1] to assert that for each property of the electron we may hope to find a corresponding property of the light quantum, and vice versa. The spin of the electron, discovered independently by Goudsmit and Uhlenbeck[2] and by Bichowsky and Urey,[3] should then have its analogue in the structure of the quantum. We feel that this idea presents difficulties, but we shall develop some of its consequences and shall discuss available data which have a bearing on it.

After the ideas which we shall present had been developed, a paper by Jordan[4] came to hand in which he accounts for the interaction of polarized light with analyzers by applying the concepts of the statistical interpretation of quantum dynamics. He determined the probability that a quantum which has passed an appropriate polarizing apparatus will be transmitted by an analyzer set at any desired angle with the polarizer. Formally, he has found it possible to associate with polarized light a "quantum-dynamical magnitude" descriptive of its state of polarization. This enables him to derive the probability that it passes the analyzer by Pauli's method for determining the probabilities that the spin vector of the electron entering a magnetic field will take a position parallel or antiparallel

to the lines of force. Granting for the moment that the polarization of the quantum and the spin of the electron are related phenomena, the exact correspondence postulated by Jordan seems to us only a special case of more general possibilities. Here we shall not adopt the language of the statistical interpretation of atomic mechanics, but shall speak of individual atoms and quanta, subject to laws which render their behavior definite. A translation of the *physical* ideas involved into the terminology of the theories advocated by Heisenberg, Dirac, and Jordan will probably not be difficult. The concepts with which we deal are the energy, momentum and angular momentum of atoms and quanta. We must examine to what extent this is allowable. Schrödinger[5] has pointed out that the concept of atomic energy may be replaced by that of frequency of the matter-waves. The interpretation of momentum in terms of the wave frequency is simple. For a single particle moving with velocity v the relation between momentum M and energy E $(= h\nu)$ is $M = h\nu\, v/c^2$. The extension to other systems is obvious. As for the interpretation of the "angular momentum of an atom," Hund's theory of spectral terms was founded on the vector properties of the quantum numbers, s, l, j, m, etc. Wigner[6] and Hund[7] have pointed out the way in which the enumeration of spectral terms can be carried out in wave mechanics without reference to angular momentum. Their conclusions are in general agreement with the original derivation of the theory and, therefore, it seems relatively unimportant whether we speak in terms of wave mechanics or of additions of angular momentum vectors. The situation is quite similar to the discussion of a branch of analysis in the language of geometry. With this understanding of the meaning of energy, momentum and angular momentum we proceed to discuss some suggested theories of light quanta.

Beck's concept of unpolarized light is an aggregate of quanta with intrinsic impulse moment vectors oriented at random, while a linearly polarized beam is analogous to a beam of electrons which have been oriented by a magnetic field. Beck makes no assumptions as to the magnitude of the impulse moment vector, in any case, except that it is not zero in linear light; nor does he specify its orientation in the various types of polarized light. In what follows we shall designate this vector by $\mathbf{r}h/2\pi$ and shall call r the *radiation* quantum number.

Our problem is to examine the values which it can assume in various types of light, consistent with known data, and to suggest experiments which may yield further information.

§2. *The Selection Principle for j.*—In the older theories the selection rules for the inner quantum number, j, could be derived by considering the conservation of angular momentum in the system, atom + field of radiation. The quantum emitted had angular momentum $h/2\pi$ or 0 according as the change in j was $\pm$ 1 or 0. In all cases the direction of j

was preserved during emission. If we assume that $r = 1$ for every quantum the same selection rule is obtained. The transitions in which j changes by unity occur just as in the older theory, but those in which j is unchanged in magnitude are accompanied by a change in its direction so that the closing side of the angular momentum triangle equals $h/2\pi$, representing the impulse moment of the light. The transition $0 \longrightarrow 0$ is impossible on this theory because a system devoid of impulse moment cannot give angular momentum to the radiation field and still have $j = 0$. This condition could not be derived in the older theory by using the correspondence principle or considerations of angular momentum, but can be derived by matrix mechanics. Of course if the q's of the matrix theory are interpreted as characteristic of the individual atom, the angular momentum changes in every emission can be calculated, resulting in a confirmation of older theory. On the other hand, the statistical interpretation of quantum mechanics means that the q's of matrix theory are related to probabilities of radiation, i.e., to average amplitudes from an aggregation of atoms. Until the interpretation of the q's is more definitely known, no decision as to the nature of the individual quanta can be definitely made. For example, linearly polarized light from an aggregate of atoms may be a superposition of linear quanta; or of circular or elliptic quanta with *appropriate phase relations*. The decision as to whether r is always one, or whether it is sometimes one and sometimes zero, is closely bound up with a decision between the macroscopic statistical interpretation of the q's and the view that they are characteristic of the individual atom.

An important fact in favor of the assumption that r is always one is that linearly polarized light can raise Hg atoms from 1^1S_0 to 2^3P_1, a transition involving a gain of $h/2\pi$ in angular momentum.[8] This is hard to explain in terms of quanta for which $r = 0$, unless we abandon the principle that $\mathbf{j} + \mathbf{r}$ is a conservative quantity.

It remains to consider the possibility that r is sometimes 2, 3, ... or perhaps $^1/_2$, $^3/_2$, ... No case is known in which $\Delta j = 2, 3, \ldots$. except in the presence of external fields. It seems possible that in such cases the integer j assigned to the terms by empirical methods may not have the significance of angular momentum. Half integral values seem to be excluded by the selection principle for j unless radical changes in the interpretation of j are introduced.

Incidentally the state of polarization of the light emitted by individual atoms has not been determined for obvious experimental reasons. An almost impossible experiment, which would contribute valuable information, is to send a beam of atoms through an inhomogeneous magnetic field and excite the resonance radiation of the Stern-Gerlach beams after they emerge from the field. A much more reasonable experiment would be to place a resonance bulb in a magnetic field strong enough to produce com-

plete orientation of the atoms (perhaps a hundred gauss). Let the field be removed very quickly. Presumably the atoms retain their orientation for a time which is very long compared to τ, the life of an excited atom, so that their resonance radiation will inform us as to the polarization of the light emitted at various angles with the j-vectors of the atoms. For example, the mean free path in a mercury resonance bulb at -20 °C. is of the order of 500 cm., which is traversed in about 10^{-3} sec. It might be thought that the limiting factor would be the field of the neighboring Hg atoms. On the average, one atom is present in every cube whose side is 1 μ. It must be remembered that experiments on the polarization of resonance radiation in the absence of a magnetic field, indicate almost complete persistence of the orientation of an excited atom. It is possible that with suitable arrangements of Kerr cells and alternating magnetic fields the experiment could be performed by continuous wave methods. A very desirable modification of this experiment would be one in which the exciting light and the magnetic field are suddenly cut off with a slight lag of the magnetic field behind the light, the resonance radiation being observed only after the removal of the field. Such an experiment would be very difficult if possible at all, but would furnish a definite test of the polarization of the light from an atom oriented by a magnetic field after its excitation.

§3. We see from the preceding section that the selection principle for j does not permit us to decide between the hypotheses, $r = 1$ and $r = 1$ or 0. We now consider emission in a magnetic field of strength suitable to produce the ordinary Zeeman effect. The older theory stated that when Δm, the change of the magnetic quantum number, is ± 1 the light has angular momentum $h/2\pi$; but when $\Delta m = 0$, the angular momentum is zero. In the first case the emitted light has the properties of a circularly polarized wave with axis parallel to the field and in the latter case it resembles the light from a linear oscillator moving parallel to the lines of force. Let us examine whether these experimental results can be explained by quanta all of which have angular momentum $h/2\pi$. Anticipating we find that this can be done by postulating a certain type of coherence between the individual quanta of a monochromatic beam.

Formally, we may consider the atom and the field-producing magnet as a single system having the quantized angular momentum $\mathbf{J}$ (in quantum units) $= \mathbf{j}_f + \mathbf{j}$, where $\mathbf{j}_f$ is the angular momentum of the field-producing mechanism, i.e., the electrons flowing in the magnet coils. j_f and j precess about J with the frequency of the Larmor precession and m is very nearly the projection of j on J because J and j_f nearly coincide in direction, both being very large compared with j. The various quantized orientations of j with respect to j_f, following the rules for coupling of vectors inside the atom, give the usual spatial orientations of j with respect to J. This

model gives us a useful understanding of the mechanism of orientation in a magnetic field. It suggests that this mechanism is the same as that which governs the coupling of the vectors belonging to individual electrons in the atom. Now that the quantization of the helium atom[9] seems to be achieved by means of the Schrödinger theory, there can be little doubt that this orientation can be described by the methods of wave mechanics applied to a model composed of one very large and one very small rotator. The problem is to follow the structure of the ψ function for such a model while the angular momentum of the large rotator grows from zero to a value very large compared to $h/2\pi$ according to any mathematically simple law of increase. It may be that the orientation can be followed by applying the relativistic wave equation to the atom alone, retaining terms of a higher order than those used in previous discussions of the way in which the Larmor precession is set up.

When $\Delta m = \pm 1$, $\Delta J = \pm 1$ and the emitted light is circularly polarized with axis along J. When $\Delta m = 0$, or $\Delta J = 0$, it is linearly polarized with the electric vector parallel to J. This model is very artificial, but if results from it can be trusted, it shows that the polarization of the light emitted by a field-free atom is in agreement with the predictions of the older theory. But, just as in §2, it may be that the linearly polarized light of the p-components is made up of light from which $r = 1$. We see that when $\Delta m = \pm 1$, the r-vector is parallel to J, i.e., very nearly parallel to H. Light of the s-components seen along the field will be left or right circularly polarized according as r is parallel or antiparallel to the direction of flight. If an s-component is observed perpendicular to the field the r-vector stands perpendicular to the direction of flight and the electric vector. In the elliptic light emitted at an angle ϑ with the field the r-vector stands at an angle ϑ to the direction of flight and in a plane perpendicular to the plane defined by the line of flight and the major axis of the ellipse. The case of the p-components is more difficult. Here the r-vectors lie in a plane perpendicular to H. Let φ be an azimuth measured in this plane and let the eye of the observer be located in the line $\varphi = 0$, looking toward the origin. Quanta having their r-vectors along the lines $\varphi = 90°$ and $\varphi = 270°$ will give linearly polarized light along $\varphi = 0$, but quanta with r-vectors at 0° and 180° appear circularly polarized when emitted along $\varphi = 0$. Intermediate orientations of the r-vector give elliptic light along $\varphi = 0$. But now we make one of two assumptions: (1) The phase relations of the circular and elliptic quanta emitted along $\varphi = 0$ are such that the resultant light is linearly polarized; (2) only the quanta with r-vector at $\varphi = 90°$ and $\varphi = 270°$ are emitted along $\varphi = 0$. Hypothesis (2) is incapable of explaining the 80% vertical linear polarization of Hg resonance radiation in the absence of a magnetic field, when the exciting electric vector is vertical.[10] Hypothesis (1) does

not seem satisfactory to the authors as compared with the simple assumption that $r = 0$ for linear light, but it is difficult to disprove on purely experimental grounds. Ideas similar to (1) have been suggested by Hanle[11] and by Breit, Ruark and Brickwedde.[12] Indeed (1) is simply a special case of the general idea that phases and amplitudes of a macroscopic radiation field have only a statistical significance. The usual formulation is that at a given point in a beam of light moving parallel to the x axis, E_y^2 and E_z^2, together with the phases of E_y and E_z, determine the number and the *resultant* polarization of the quanta arriving at that point.

With the aid of hypothesis (1) and using quanta for which $r = 1$, it is possible to explain all the cases of polarization of resonance radiation which have yielded to previous theories as well as the polarization of light emitted in an electric field. In discussing resonance radiation in a magnetic field we meet with cases where the exciting light has a polarization different from that which would be emitted by the vapor toward the source of the exciting light if excited by electronic impact while in the field. For example, Hanle has illuminated a mercury resonance bulb by both circular and elliptic light. When no field is present the character of the light emitted at a small angle (20°) with the incident beam is approximately the same as that of the exciting light. On applying a field parallel to the incident beam, circularly polarized light is unaffected; but if the incident light be elliptically polarized, the resonance radiation contains circularly polarized light and also depolarized light. Evidently some of the atoms in the resonance bulb have been raised to the Zeeman level $m = 1$ and re-emit circularly polarized light. The fate of the others which give rise to the depolarized light is uncertain; perhaps they are re-oriented after absorption to the positions $m = 0$ and $m = -1$. The character of the light emitted by these atoms lies beyond our present knowledge, and probably will not be amenable to theory until we understand the way in which atoms arrive at their quantized orientations when a magnetic field is applied.

The theory of the spinning quantum may be applied to aid in the interpretation of experiments on the polarization of radiation excited by unidirectional electron impacts. When λ 2536.7 of Hg is excited in this way, the radiation is partially polarized, with the electric vector perpendicular to the exciting beam of electrons. This means that the r-vectors of the quanta lie parallel and antiparallel to the electron beam. This is in some measure understandable if we suppose that the impulse moment vectors of the electrons lie in the direction of the accelerating field, being space-quantized in analogy to the orientation of hydrogenic atoms in the Stark effect. Of course, many other factors are involved in the complete explanation of this experiment. Skinner[13] has already suggested that the spin of the electron plays an important part in experiments of

this kind. Further experiments in which atoms are bombarded by electrons while exposed to a magnetic field parallel to the electron beam are much to be desired.

§4.—It is apparent from what precedes that if r is always 1 there may be several kinds of linearly polarized light. In the normal Zeeman effect, the s-component of higher frequency is composed of quanta with r-vectors parallel to H, while that of lower frequency has r-vectors antiparallel to H. We shall refer to these two types when observed transversely as positive and negative linear quanta, respectively. If the origin of such quanta is unknown they may be called right or left handed, according as the r-vector points to the right or left of the observer. A mixture of these two types may be called racemic light composed of linear quanta. Further we may have linear light composed of circular or elliptic quanta; and, finally, if r is sometimes zero, we may have a "zero" variety of linear light. Similarly, there may be several varieties of circular or elliptic light. Whenever polarized light is produced from unpolarized light by gross polarizing apparatus, we probably deal with racemic varieties. *However, the light composing the various lines of a Zeeman pattern is generally non-racemic.* Thus the Zeeman pattern is to be regarded as a sort of Stern-Gerlach pattern for quanta. We wish to suggest several experiments to test the validity of these views.

A.—Let the light from a mercury resonance bulb placed in a strong magnetic field shine on a second resonance bulb in which the earth's field is very carefully neutralized. Only the p-component will have the frequency necessary to excite the second bulb. If r is zero for this component, there should be no absorption, but if r is 1, absorption is possible. It must be understood in this as in the succeeding experiments that we are speaking of an idealized case in which the line has no fine structure, the Zeeman effect being the $^3/_2$-normal triplet. In practice difficulties will arise because this line is composed of five fine structure components, some of which behave abnormally in the magnetic field.[10]

B.—Consider two resonance bulbs lying on the same east and west line and in the same horizontal plane, in magnetic fields of identical strength. Let the field on the first bulb be directed to the north while that on the second can be placed at any azimuth in the horizontal plane. Initially let it point north. Let the first bulb be illuminated by a mercury arc. Light from the first bulb falling on the second should be absorbed in a quite ordinary way and the emitted light should have the same character as that from the first. Now let the second field point south. A quantum of the s-component of higher frequency will leave the first bulb with r-vector pointing north. There are atoms in the second bulb which have suitable energy levels for absorbing this frequency, but in absorbing it they would acquire angular momentum from the light directed toward

the north, whereas the transition to the upper Zeeman level would require an increase of angular momentum directed to the south. Similar relations hold for the *s*-component of lower frequency. Since the angular momentum vector for the field-producing mechanism (j_f of §3), does not change in the emission process, it appears reasonable that it cannot change in absorption, so that the angular momentum cannot be brought to its normal value by a contribution from the field. Therefore, an absorption of the kind described would be impossible if angular momentum is to be conserved in each individual process.

C.—Place the resonance bulbs of the previous experiment on the same north and south line. Let the field on one point north and that on the other point south. The light of the *s*-component of higher frequency emitted by the first bulb will have its *r*-vectors pointing north. But if atoms in the second resonance bulb are carried to the upper Zeeman level of 2^3P_1 they should receive a quantum with angular momentum vector pointing south. Similar relations hold for the *s*-component of lower frequency. Again there is no reason to believe that this angular momentum would be supplied by the field, so that absorption should not occur. Indeed, it is known from experiments on the Zeeman effect in resonance radiation that the wrong sign of circularly polarized light will not be absorbed.[14] We have included this experiment simply because it has not been performed with a resonance lamp as the source of the polarized light and there may be some doubt as to what would happen in this case.

D.—It may be interesting to consider how non-racemic monochromatic linear light (or, for that matter, non-racemic circular or elliptic light) can be obtained. Let a resonance bulb in a strong magnetic field be illuminated by a source which coincides with only one of the *s*-components. If the element in the resonance bulb is properly chosen, e. g., Hg, there will be only one component of the Zeeman pattern in the resonance light. In other cases there will be several lines in the emitted light such as those discussed by Foote, Ruark and Mohler.[15] Experiments in which a single Zeeman component is used as a source have an interest far wider than the proof or disproof of the theories here presented. A typical experiment of this kind bearing on the theory of the spinning quantum is as follows: Take right-handed axes with OY vertical. Let a linearly polarized quantum of one of the *s*-components of $1^1S_0-2^3P_1$ fall on a resonance bulb in the direction OX, the electric vector being vertical, and the *r*-vector along OZ. Will the resonance radiation along OZ be partly or wholly circularly polarized?

The authors wish to emphasize that this paper is not a defense of any particular theory. It aims to show that the theory of spinning quanta can explain a great variety of experimental results and to suggest further experiments on which it must stand or fall.

[1] *Z. Physik*, **43**, 658 (1927).
[2] *Naturwissenschaften*, Nov. 20, **1925**.
[3] *Proc. Nat. Acad. Sci.*, Feb., **1926**.
[4] *Z. Physik*, **44**, 292 (1927).
[5] *Ann. Physik*, **83**, 956 (1927).
[6] *Z. Physik*, **40**, 492; **43**, 624 (1927).
[7] *Ibid.*, **43**, 788 (1927).
[8] Eldridge, *Phys. Rev.*, **24**, 234 (1924).
[9] G. W. Kellner, *Z. Physik*, **44**, 91 (1927).
[10] Ellett and Macnair, *Proc. Nat. Acad. Sci.*, Aug., **1927**. The departure from 100% is due to the fine structure components which behave anomalously.
[11] *Z. Physik*, **30**, 93 (1924).
[12] *Phil. Mag.*, **3**, 1306 (1927).
[13] *Proc. Roy. Soc.*, **112A**, 642 (1926).
[14] W. Hanle, *Z. Physik*, **41**, 164 (1927).
[15] *J. O. S. A. and R. S. I.*, **7**, 415 (1923).

THE SYMMETRY OF THE STRESS-TENSOR OBTAINED BY SCHROEDINGER'S RULE

BY H. BATEMAN

CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated October 3, 1927

In the recent developments of the calculus of variations used in the new quantum theory the problem arises of finding a general expression for a world-function such that a symmetrical stress-energy tensor may be derived from it by means of Schroedinger's rule.[1]

We shall consider here the analogous problem for the case of a Euclidean space of three dimensions in which the rectangular coördinates of a selected point P are x, y, z. To simplify matters we shall consider a world-function L which depends only on the first derivatives of the components u, v, w of a single vector q associated with the point P. Applying the rule used by Schroedinger in his discussion of Gordon's equations, but with the necessary modifications appropriate for a space of three dimensions, we may associate with L a tensor T with mixed components[2] of type

$$T_{23} = u_2 \frac{\partial L}{\partial u_3} + v_2 \frac{\partial L}{\partial v_3} + w_2 \frac{\partial L}{\partial w_3} + v_1 \frac{\partial L}{\partial w_1} + v_2 \frac{\partial L}{\partial w_2} + v_3 \frac{\partial L}{\partial w_3}$$

the suffixes 1, 2, 3 being used to denote differentiations with respect to x, y, z, respectively.

In order that this tensor may be symmetric the relations

$$T_{23} = T_{32}, \qquad T_{31} = T_{13}, \qquad T_{12} = T_{21}$$

must be satisfied and these give the partial differential equations

$$(v_2 - w_3)\left(\frac{\partial L}{\partial v_3} + \frac{\partial L}{\partial w_2}\right) + (v_3 + w_2)\left(\frac{\partial L}{\partial w_3} - \frac{\partial L}{\partial v_2}\right) + v_1\frac{\partial L}{\partial w_1} - w_1\frac{\partial L}{\partial v_1} + u_2\frac{\partial L}{\partial u_3} - u_3\frac{\partial L}{\partial u_2} = 0$$

$$(w_3 - u_1)\left(\frac{\partial L}{\partial w_1} + \frac{\partial L}{\partial u_3}\right) + (w_1 + u_3)\left(\frac{\partial L}{\partial u_1} - \frac{\partial L}{\partial w_3}\right) + w_2\frac{\partial L}{\partial u_2} - u_2\frac{\partial L}{\partial w_2} + v_3\frac{\partial L}{\partial v_1} - v_1\frac{\partial L}{\partial v_3} = 0$$

$$(u_1 - v_2)\left(\frac{\partial L}{\partial v_1} + \frac{\partial L}{\partial u_2}\right) + (u_2 + v_1)\left(\frac{\partial L}{\partial v_2} - \frac{\partial L}{\partial u_1}\right) + u_3\frac{\partial L}{\partial v_3} - v_3\frac{\partial L}{\partial u_3} + w_1\frac{\partial L}{\partial w_2} - w_2\frac{\partial L}{\partial w_1} = 0.$$

These equations form a complete system and since there are nine independent variables and three equations, the general solution should be of the form

$$L = f(\alpha, \beta, \gamma, \theta, \phi, \psi)$$

where $\alpha, \beta, \gamma, \theta, \phi, \psi$ are independent solutions.

Writing

$$a = u_1, \qquad b = v_2, \qquad c = w_3,$$

$$2f = w_2 + v_3, \quad 2g = u_3 + w_1, \quad 2h = v_1 + u_2,$$
$$2\xi = w_2 - v_3, \quad 2\eta = u_3 - w, \quad 2\zeta = v_1 - u_2$$

we choose as our independent solutions the quantities

$$\alpha = \xi^2 + \eta^2 + \zeta^2$$
$$\beta = a\xi^2 + b\eta^2 + c\zeta^2 + 2f\eta\zeta + 2g\zeta\xi + 2h\xi\eta$$
$$\gamma = (bc - f^2)\xi^2 + (ca - g^2)\eta^2 + (ab - h^2)\zeta^2 + 2(gh - af)\eta\zeta + 2(hf - bg)\zeta\xi + 2(fg - ch)\xi\eta$$

$$\gamma = a + b + c$$
$$\phi = bc - f^2 + ca - g^2 + ab - h^2$$
$$\psi = abc + 2fgh - af^2 - bg^2 - ch^2$$

If

$$A = u_1^2 + v_1^2 + w_1^2 \qquad F = u_2u_3 + v_2v_3 + w_2w_3$$
$$B = u_2^2 + v_2^2 + w_2^2 \qquad G = u_3u_1 + v_3v_1 + w_3w_1$$
$$C = u_3^2 + v_3^2 + w_3^2 \qquad H = u_1u_2 + v_1v_2 + w_1w_2$$

it is easily seen that the following quantities are also solutions

$$\rho = A + B + C$$

$$\sigma = BC - F^2 + CA - G^2 + AB - H^2$$
$$\tau = ABC + 2FGH - AF^2 - BG^2 - CH^2$$
$$\omega = A\xi^2 + B\eta^2 + C\zeta^2 + 2F\eta\zeta + 2G\zeta\xi + 2H\xi\eta$$

$$\Delta = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

$$\mathrm{X} = v_2w_3 - v_3w_2 + w_3u_1 - w_1u_3 + u_1v_2 - u_2v_1$$

These quantities can, however, be expressed in terms of α, β, γ, θ, ϕ, ψ for we have the relations

$$2\rho = \alpha + 2\theta^2 + 4\phi$$
$$\sigma = (\alpha - \phi)^2 + 4\gamma + 2\theta(\beta - \psi)$$
$$\tau = \Delta^2$$
$$\omega = \gamma + \theta\beta - \phi\alpha$$
$$4\Delta = \beta + 4\psi$$
$$\mathrm{X} = \phi + \alpha.$$

Many of the solutions which have been written down represent quantities of importance in the theory of finite displacements and the theory of elasticity.[3] The three-dimensional problem was, in fact, studied with the aim of finding a general type of world function which can be used in the theory of elasticity. For the purpose of extending the present results to a space of n dimensions and in particular to space-time geometry it will be useful to note a very general type of solution of our partial differential equations. Let p, q, r, s be arbitrary parameters then the partial differential equations are satisfied by the determinant

$$L = \begin{vmatrix} pu_1 + qa + rA + s & pu_2 + qh + rH & pu_3 + qg + rG \\ pv_1 + qh + rH & pv_2 + qb + rB + s & pv_3 + qf + rF \\ pw_1 + qg + rG & pw_2 + qf + rF & pw_3 + qc + rC + s \end{vmatrix}$$

Another point which may be mentioned is that the differential equations are unaltered in form if we replace u, v, w by u', v', w' where

$$u' = u + kx, \qquad v' = v + ky, \qquad w' = w + kz$$

and k is an arbitrary parameter. Using primes to denote quantities derived from u', v', w', we have

$$u_1^1 = u_1 + k \qquad u_2' = u_2 \qquad u_3' = u_3$$
$$v_1^1 = v_1 \qquad v_2' = v_2 + k \qquad v_3' = v_3$$
$$w_1^1 = w_1 \qquad w_2' = w_2 \qquad w_3' = w_3 + k$$
$$\xi^1 = \xi \qquad \eta^1 = \eta \qquad \zeta^1 = \zeta$$
$$\alpha^1 = \alpha \qquad \beta^1 = \beta + k\alpha \qquad \gamma^1 = \gamma + k(\alpha\theta - \beta) + k^2\alpha$$
$$\theta^1 = \theta + 3k \qquad \phi^1 = \phi + 2k\theta + k^2 \qquad \psi^1 = \psi + k\phi + k^2\theta + k^3$$

It is clear that if any solution of our equations is expressed in terms of the derivatives of u', v', w' and then expanded in powers of k the coefficients of the different powers of k will all be solutions of the equations.

[1] E. Schroedinger, *Ann. Physik*, **82** (1927), 265.

[2] We call T_{11} a pure and T_{23} a mixed component. The expression for T_{11} is found by a somewhat similar rule except that there is an additional term—L.

[3] See, for instance, Love's Treatise, Ch. I, Appendix. L. Brillouin, *Ann. Physique*, (10) **3** (1925), 251.

SOME EXTENSIONS OF THEORY AND MEASUREMENTS OF SHOT-EFFECT IN PERIODIC CIRCUITS

By H. B. Vincent*

Cruft Laboratory, Harvard University

Communicated October 14, 1927

Introduction.—The paper published last year by Williams and Vincent[1] contained computations of the expected magnitude of the effect in an aperiodic circuit of the probability fluctuations of anode current in a vacuum tube. The measurements submitted were made with a system of this type. There appeared there also a discussion of the problem of the tuned circuit in which a method was suggested to simplify the necessary measurements. Since this paper is concerned with some additions to this method and measurements made thereon, a brief résumé seems desirable here.

Simplification of the Tuned Circuit Problem.—The power factor of the circuit varies with the thermionic current, but it may be maintained constant by the addition of suitable resistances. Let this be done.

If now we take the expression given by Hull and Williams[2] for the mean square output voltage due to the k-component of the thermionic current, we have

$$\overline{E_k^2} = {}^1/_2 L^2 \omega_0^2 A_0^2 C_k^2 \frac{x^2 f(x)}{(1 - x^2)^2 + r^2 x^2},$$

where

L is the inductance in the shot-circuit,
ω_0 is 2π times the resonance frequency of the amplifier,
A_0 is the voltage amplification at the resonance frequency,
C_k is the coefficient of the k-component of the Fourier series into which the thermionic current was originally developed by Schottky[3]
$x = \omega/\omega_0$

$f(x)$ is the curve obtained by plotting relative voltage square amplification against x,

$r = R/L\omega_0$, where R is the equivalent series resistance of the shot-circuit.

But[4]

$$\overline{E_k^2} = {}^1/_2 L^2\omega_0^2 A_0^2 C_k^2 \frac{x^2}{(1-x^2)^2 + r^2x^2} \times \frac{v_{0c}^2}{v_{kc}^2}\varphi(x),$$

where $\varphi(x)$ is the curve obtained when the input current and not the condenser voltage is held constant and x varied. Now[4]

$$\frac{v_{0c}^2}{v_{kc}^2} = \frac{\omega_k^2 + (L^2/R^2)(\omega_k^2 - \omega_0^2)^2}{\omega_k^2} = \frac{(1-x^2)^2 + r^2x^2}{r^2x^2}.$$

So that

$$\overline{E_k^2} = \frac{L^2A_0^2}{2R^2C^2} C_k^2\varphi(x)$$

where C is the value of the capacitance in the shot-circuit.

Substituting for C_k^2 its value[3]

$$\overline{C_k^2} = 4I_0\epsilon\omega_0 dx/2\pi,$$

where ϵ is the value of electronic charge and I_0 is the average value of the thermionic current traversing the shot-circuit, we obtain for the total mean square output voltage

$$\overline{E^2} = \Sigma\overline{E_k^2} = \frac{L^2A_0^2}{2R^2C^2} \times \frac{4I_0\epsilon\omega_0}{2\pi} \sum\varphi(x)dx = \frac{2I_0\epsilon L^2A_0^2}{R^2C^2}\int_0^\infty \psi(n)dn,$$

where n (the frequency) $= \omega/2\pi$, $\psi(n) = \varphi(x)$, and $dx = d(\omega/\omega_0) = (1/\omega_0)d(2\pi n) = (2\pi/\omega_0)dn$. $\int\psi(n)dn$ becomes simply the area under the curve $\varphi(x)$ plotted against frequency. If we call this area A, the above equation yields

$$E^2 = 2I_0\epsilon L^2A_0^2A/R^2C^2.$$

If now we introduce a voltage v_1 which causes the same output as the shot-effect, we have

$$I = v_1/R,$$

where I is the current in the tuned circuit and v_1 has the frequency $n_0 = \omega_0/2\pi$, and the condenser voltage v_c is

$$v_c = I/\omega_0 C = v_1/RC\omega_0.$$

Also

$$E^2 = A_0^2v_c^2 = A_0^2v_1^2/R^2C^2\omega_0^2 = 2I_0\epsilon L^2A_0^2A/R^2C^2.$$

So that

$$\epsilon = v_1^2/2L^2\omega_0^2I_0A.$$

It may be noted that by this means we may avoid the complicated evaluations of the original shot-integral. Also if v_1 is introduced by a current I_1 flowing through an inductance L_1 of negligible resistance, we have

$$\epsilon = \frac{L_1^2 I_1^2 \omega_0^2}{2L^2 I_0 A \omega_0^2} = \frac{L_1^2 I_1^2}{2L^2 I_0 A}.$$

This measurement is seen to be independent of all frequency measurements except those of frequency differences for the area A.

Additional Simplifications.—In the following discussion G_1 will be used to denote readings of the galvanometer which indicates the current from the generator and G_2 to denote readings in the amplifier output.

Let now a current I_0 be sent through the tube and the resulting deflection in the output noted. Call this reading G_{21}. The generator may now be started and the current adjusted to give some suitable reading on the galvanometer G_1. Call this reading G_{11}. The reading of G_2 will now be increased and this increase will be the ordinate of a curve $F(n)$ for the frequency, n, of the generator. The generator frequency may be varied with G_{11} held constant and the curve plotted. The area under this curve divided by its maximum ordinate yields the area A of the preceding section.

The thermionic current traverses the circuit at all times so that no question arises concerning the constancy of the power factor of the circuit. No error is introduced by the addition of a resistance whose value may be slightly incorrect. No new assumption is involved since we have previously assumed the addition of the energies of the various components present.

If we denote the area under the curve $F(n)$ by A_1 and the maximum output deflection by G_{2m}, the expression for ϵ is as follows:

$$\epsilon = \frac{L_1^2 I_1^2}{2L^2 I_0 \dfrac{A_1}{G_{2m} - G_{21}}},$$

and this may be further simplified in case the deflection of the input galvanometer is proportional to the square of the input current. Let I denote the input current. Then

$$KG_1 = I^2, \quad KG_{11} = I'^2 \text{ where } K \text{ is a constant.}$$

But since the output deflection must be proportional to the input voltage squared,

$$G_{2m} - G_{21} = kI'^2 = kKG_{11},$$

where k is a constant.

Now the particular I_1 required by the equation is a current such that

$$G_{2m} - G_{21} = G_{21},$$

and

$$G_{21} = \frac{G_{21}kKG_{11}}{G_{2m} - G_{21}} = kI_1^2, \quad \text{so that} \quad I_1^2 = \frac{KG_{11}G_{21}}{G_{2m} - G_{21}}$$

and the equation now becomes

$$\epsilon = \frac{KL_1^2G_{11}G_{21}}{2L^2I_0A_1}.$$

K is determined in the calibration of the system measuring the input current—a vacuum thermocouple and galvanometer. No new quantities are introduced which did not before enter into the measurements and one, the maximum ordinate, is eliminated.

Mechanical Integration of Area under Resonance Curve.—With the system described above a curve $F(n)$ is to be taken for every value of ϵ obtained. This gives an excellent independence of results, but considerable time is required for obtaining the data, plotting the curve and integrating. Further, the action of the amplifier must be the same at all times while data for $F(n)$ is being taken. When the amplification used is of the order of a million fold at a frequency of 10^6 cycles per second, this places a requirement of great constancy on the batteries actuating the system. Thus it would seem to be a considerable advance in both speed and accuracy if a method could be devised by which the comparison of voltages might be made in a very brief interval of time. Such a method has been developed and is presented herewith.

The system employs an oscillator designed to give a current of suitable frequency and as free from harmonics as possible. A small variable condenser of the rotating plate type is connected in parallel with the principal tuning capacity. We may refer to this condenser as C_1.

C_1 is to be designed to give straight-line capacity. That is, if we reckon displacement from the position of minimum capacity, the capacity increase is to be proportional to the angular displacement in either sense of rotation up to the maximum value. In actual practice this condition does not hold for small angles, but the difficulties involved may be eliminated as will be explained later. For the mathematical discussion we shall assume a strict proportionality.

Let us arrange conditions so that the maximum and minimum frequencies, generated as C_1 is rotated, differ by the same amount from the resonance frequency n_1 of the amplifier. If we choose the condenser position where the resonance frequency is generated as our zero, if the capacity increase per radian of C_1 is b, and if we designate the angular displacement from this zero by θ, we have

$$C_1 = b\theta \text{ for } -\pi/2 \leqq \theta \leqq \pi/2.$$

Let us call the tuning capacity of the oscillatory circuit C_2 for the zero position of C_1. Then the angular velocity of the wave developed for any position of C_1 is given by the expression

$$\omega = \frac{1}{\sqrt{L(C_1 + C_2)}} = \frac{1}{\sqrt{LC_2}}\left(1 + \frac{C_1}{C_2}\right)^{-1/2}$$

and if C_1 is small compared to C_2 so that squares and higher powers of C_1/C_2 may be neglected,

$$\omega = \omega_0\left(1 - \frac{C_1}{2C_2}\right) = \omega_0\left(1 - \frac{b\theta}{2C_2}\right).$$

The condenser C_1 is connected to a motor and driven with a constant angular velocity ω_1, so that

$$\theta = \omega_1 t.$$

If the oscillator be so designed that the amplitude of the output current is the same for all values of C_1, this current, I, will be of the form

$$I = a \cos \omega t = a \cos \omega_0\left(1 - \frac{b\omega_1 t}{2C_2}\right) t,$$

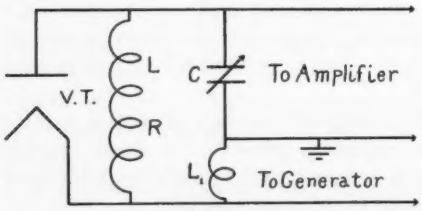


FIGURE 1

Diagram showing short-circuit and method of introduction of calibrating voltage.

where a is the amplitude.

This is the current which will traverse the inductance L_1, figure 1, the part of the current passing through the tuned circuit being negligible in comparison with the whole.

The discussion may now be carried on most briefly in terms of complex angles where

$$I = \text{real part of } ae^{j\omega t}.$$

The voltage v_1 across L_1 is

$$v_1 = jL_1\omega a e^{j\omega t}$$

and the series resistance of the tuned circuit is

$$Z_s = R + j(L\omega - 1/C\omega).$$

If the amplitude of the current in the tuned circuit is I_c and the condenser voltage v_c, we have

$$I_c = v_1/Z_s$$

and

$$v_c = I_c/jC\omega = v_1/jC\omega Z_s$$

$$= \frac{jL_2\omega a e^{j\omega t}}{jC\omega[R + j(L\omega - 1/C\omega)]}$$

$$= \frac{L_1 a e^{j\omega t}}{C[R + j(L\omega - 1/C\omega)]}$$

and the real part of the condenser voltage

$$v_c \cos \omega t = \frac{L_1 a}{C\sqrt{R^2 + (L\omega - 1/C\omega)^2}} \cos (\omega - \varphi)t$$

where $\varphi = \tan^{-1} (L\omega - 1/C\omega)/R$.

This angle of retardation φ is negligible in comparison to ω for narrow bands of high frequency, the maximum value of φ/ω amounting in the apparatus used to about 10^{-6}.

Now the output deflection of the amplifier which will be denoted by G_2' is proportional to the square of the output voltage. Thus we have

$$E^2 = BG_2' = A_0^2 f(x)\overline{v_c^2}$$

where A_0 and $f(x)$ have the significance previously assigned and B is a factor of proportionality.

Thus

$$BG_2' = \frac{1}{t}\int_0^t [A_0\sqrt{f(x)}v_c \cos \omega t]^2 dt$$

$$= \frac{A_0^2}{t}\int_0^t f(x)v_c^2 \cos^2 \omega t \, dt$$

and for the particular period in which we are interested

$$BG_2' = A_0^2 \frac{\omega_1}{\pi}\int_{-\pi/2\omega_1}^{\pi/2\omega_1} f(x)\left[\frac{L_1 a}{C\sqrt{R^2 + (L\omega - 1/C\omega)^2}}\right]^2 \cos^2 \omega t \, dt.$$

$$= \frac{A_0^2 L_1^2 a^2}{C^2} \times \frac{\omega_1}{\pi}\int_{-\pi/2\omega_1}^{\pi/2\omega_1} \frac{f(x)}{R^2 + (L\omega - 1/C\omega)^2} \cos^2 \omega t \, dt.$$

But we have shown above that

$$\omega = \omega_0(1 - b\omega_1 t/2C_2).$$

So that

$$\omega/\omega_0 = x = 1 - b\omega_1 t/2C_2$$

and

$$b\omega_1 t/2C_2 = 1 - x, \qquad t = (1 - x)2C_2/b\omega_1, \qquad dt = -2C_2 dx/b\omega_1$$

Also for $t = -\pi/2\omega_1$, $x = 1 + b\pi/4C_2$

and for $t = \pi/2\omega_1$, $x = 1 - b\pi/4C_2$

so that

$$BG_2' = \frac{A_0^2L_1^2a^2}{C^2} \times \frac{\omega_1}{\pi} \int_{1-b\pi/4C_2}^{1+b\pi/4C_2} \frac{f(x)}{R^2 + (L\omega - 1/C\omega)^2} \times \frac{2C_2}{b\omega_1} \cos^2(2C_2/b\omega_1)(1-x)\omega\, dx$$

$$= \frac{A_0^2L_1^2a^2}{2L^2} \times \frac{2C_2}{b\pi} \left[\int_{1+b\pi/4C_2}^{1-b\pi/4C_2} \frac{L^2\omega^2 f(x)}{R^2C^2\omega^2 + (1 - LC\omega^2)^2}\, dx \right.$$

$$\left. + \int_{1-b\pi/4C_2}^{1+b\pi/4C_2} \frac{L^2\omega^2 f(x)}{R^2C^2\omega^2 + (1 - LC\omega^2)^2} \cos\frac{4C_2}{b\omega_1}(1 - x)\omega\, dx \right].$$

If now the region of integration, $1 + b\pi/4C_2$ to $1-b\pi/4C_2$, includes all values of x for which $f(x) \neq 0$, then the second integral on the right vanishes and we have

$$BG_2' = \frac{A_0^2L_1^2a^2}{2L^2} \times \frac{2C_2}{b\pi} \int_{1+b\pi/4C_2}^{1-b\pi/4C_2} \frac{L^2\omega^2 f(x)}{R^2C^2\omega^2 + (1 - LC\omega^2)^2}\, dx \quad \ldots \quad (1)$$

Let us turn our attention to the form of the expression for the output occasioned by the shot-effect. Taking the k-component quoted above

$$\overline{E_k^2} = {}^1/_2L^2\omega_0^2A_0^2\overline{C_k^2} \frac{x^2 f(x)}{(1 - x^2)^2 + r^2x^2}$$

$$= {}^1/_2A_0^2\overline{C_k^2} \frac{L^2\omega^2 f(x)}{R^2C^2\omega^2 + (1 - LC\omega^2)^2}$$

we may perform a summation on substituting for C_k^2 and obtain

$$\overline{E^2} = 1/\pi(A_0^2I_0\epsilon\omega_0) \int_{1+\kappa}^{1-\kappa} \frac{L^2\omega^2 f(x)}{R^2C^2\omega^2 + (1 - LC\omega^2)^2} \quad \ldots\ldots \quad (2)$$

where the interval $1-\kappa$ to $1 + \kappa$ includes the whole region for which $f(x) \neq 0$.

If the output deflection corresponding to this output voltage squared be denoted by G_{21}, we have

$$\overline{E^2} = BG_{21}$$

where B has the same significance and value as in equation (1).

The integrals in equations (1) and (2) refer to the same circuit and if the power factor of the circuit is the same for the two inputs, these inte-

grals will be the same. Let this condition be maintained. We have then

$$\frac{G_2'}{G_{21}} \frac{a^2 L_1^2 \pi}{2L^2 I_0 \epsilon \omega_0} \times \frac{2C_2}{b\pi}.$$

If x_1 and x_2 denote the limits between which x varies,

$$\frac{2C_2}{b\pi} = \frac{1}{x_1 - x_2} = \frac{\omega_0}{2\pi(n_1 - n_2)},$$

where n_1 and n_2 are corresponding frequencies.

Thus

$$\frac{G_2'}{G_{21}} = \frac{a^2}{2} \times \frac{L_1^2}{2L^2 I_0 \epsilon (n_1 - n_2)}.$$

Now $a^2/2$ is the mean square current from the generator which is, as in the preceding section, proportional to the deflection of the input galvanometer, so that

$$a^2/2 = KG_{11}.$$

In order that the condition of constancy of power factor be maintained the procedure may be as in the preceding section. The shot-effect is used as input and the deflection G_{21} noted, next the shot-effect and generator current are the input and the deflection is G_{22}.

Then

$$G_2' = G_{22} - G_{21}.$$

Substituting and solving for ϵ,

$$\epsilon = \frac{L_1^2 K G_{11} G_{21}}{2L^2 I_0 (n_1 - n_2)(G_{22} - G_{21})}.$$

Thus for a particular current and circuit, the value of ϵ is given in terms of three galvanometer deflections and a predetermined constant. Comparisons may be made rapidly and it is not required that the apparatus should act in precisely the same way over long periods of time.

The conditions quoted hold only for $-\pi/2 \leqq \theta \leqq \pi/2$, but the integration is obviously performed continuously in alternate senses as the condenser rotates.

It was mentioned above that the capacity of C_1 does not need to be straight line over its whole scale. It is now obvious that if the frequency generated is straight line over the whole region for which $f(x) \neq 0$, the condition is sufficient. The interval $n_1 - n_2$ is now determined by extending this straight line portion of the curve rather than from the differences of actual maximum and minimum frequencies developed.

Apparatus.—The amplifier used for this work is of the same type as that

used last year with an improvement in shielding to acquire efficiency at higher frequencies. Papers by Hull and Williams[2] have contained complete descriptions of the system. The screen-grid tubes were supplied through the courtesy of Dr. Hull of the Research Laboratories of the General Electric Company.

The detection system is of the same type as that described in the paper by Williams and Vincent[1] as is the inductance L_1, with the exception that a correction for penetration formerly neglected is here added. The oscillator is of the tuned-plate-circuit type, being shielded with sheet copper. The shot-circuit is as shown in figure 1. With this type of circuit at a frequency of 10^6 cycles per second, inductive coupling between it and the input leads from the oscillator was at first very troublesome. The elimination of this coupling was accomplished by a suitable design of shielding which must be made of metal sheet and strictly continuous throughout so that no current may appear on its outer surface. The inductance L_1 and, more particularly, the lead from it to the shot-circuit are shielded by a six-inch brass tube, one-half of which is movable for adjustments. It was found necessary to use only one-half of the heater element in the thermocouple measuring the generator input in order that the galvanometer connected thereto might be maintained at approximately the same potential as the shielding.

The tubes used to produce the shot-effect were principally U X 201A, with an anode potential of 200 volts. All tubes used yielded sensibly the same results.

Frequency measurements are effected by means of a wave-meter built into the oscillatory system so that the coupling may be invariant. Calibration was made by means of harmonics of the order of 30 of a crystal oscillator whose fundamental is known to $\pm$ 1 cycle per second.

Calibrations.—The wave-meter condenser is a General Radio Precision Condenser of straight line capacity, so that λ^2 plotted against C^0 should give a straight line where λ is the wave-length and C^0 the reading of the condenser scale. The slope of this line was determined by the method of least squares as was also the limit of error in this slope, the limit of error yielded being 0.07%.

The square of the generator current plotted against the deflection of the galvanometer connected to the thermocouple carrying this current gives a straight line with considerable precision. Many calibrations of this were carried out, a typical set yielding by the method of least squares a limit of error of 0.11% in the slope of the line. The sensitivity of the particular thermocouple used—Rawson, 750 ohms, changed continuously with use. This necessitated constant calibration, and in every case the system was linear within the limits of error of the individual readings.

The deflection of the galvanometer in the output of the amplifier should

be proportional to the deflection of the galvanometer in the generator input. This condition holds within the limits of error of reading of the galvanometers.

The inductance L_1 was constructed with great care in the machine shop at Cruft Laboratory. Dimensions are known to $\frac{1}{1000}$ of an inch and the ratios of deflections from the various taps check excellently well.

The rotating condenser, C_1, consists of a single rotor plate and two stator plates. A $\frac{5}{8}$-inch shaft running in ball bearings carries the rotor and a special graduated dial for calibration, the whole system being balanced to avoid vibration. The bearings are adjustable so that endwise motion of the shaft may be eliminated. Connection to the rotor is effected by means of a phosphor bronze brush bearing on one end of the shaft. The other end of the shaft is connected by means of a heavy-walled rubber tube to a d. c. motor whose speed may be controlled. The whole system is supported on felt to minimize vibration.

Measurements.—Making very generous allowance for possible errors in the standards used in calibration, etc., it seems impossible to account for a total error or more than two per cent. On the other hand, the values obtained differ from the accepted value of electronic charge by more than this amount. For this reason more time was spent on the investigation of possible sources of error than in obtaining values of ϵ. Several hundred values were obtained, however—all of them too large. Following the last adjustment—that of using only one-half of the heater element in the input thermocouple, the average of the values obtained is very nearly 1.7×10^{-19} coulombs. It seems unnecessary to quote a long list of individual values of this order of magnitude.

Considerable gratification is felt in the fact that values obtained by the use of the rotating condenser checked very closely those obtained by other methods, the results being on the whole more consistent.

The following is perhaps of interest. The equivalent series resistance of the shot-circuit increases with the thermionic current. We may neglect this increase and proceed as indicated under "Simplification of the Tuned Circuit Problem," assuming that the power factor does not change. This should yield values which are too small, but measurements made in this way give values still too large.

These high values may result from errors in the system which the tests employed have not revealed. However, it seems impossible to see where anything has remained unchecked. If no error is involved, the increase in the apparent value of ϵ is perhaps due to secondary emission from the anode. These electrons, returning to the anode after a very brief interval, would produce components of very high frequency which would have little effect on measurements made at lower frequencies.

Effect of High Current Densities.—Previous work on shot-effect has made mention of the decrease of observed fluctuations with space charge.[5] In this work the grid was placed at a constant negative potential with respect to the cathode, the anode potential being also constant. The space charge was increased by increasing the filament temperature and it was found that the ratio of observed to calculated mean square voltage became less with rising temperature. This diminution was supposed to be caused by the existence of the space charge but it seemed that in these experiments two changes were made simultaneously—the increase of space charge and

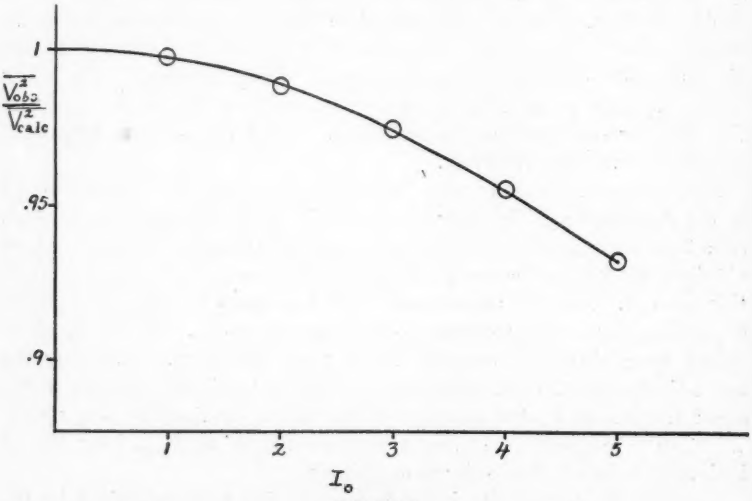


FIGURE 2

Diagram showing plot of $\overline{V^2_{obs.}}/\overline{V^2_{calc.}}$ against I_0 in milliamperes together with points of the curve given below.

an increase of anode current. Last year[1] measurements were made on the effect of space charge in the absence of any change in anode current. Very little change was observed over the whole range possible for the tube, the diminution found (about 2%) being within the limit of error of the measurements.

Observations have been made here to determine the effect on the observed fluctuations of high current densities. The plate and grid potential used being 200 volts, it would seem that there should be very little space charge for even fairly large currents. The graph shown in figure 2 is plotted from some 70 measurements.

An attempt was made to find a mathematical expression for this curve

and the result is shown by the points marked thus ⊙. These are points of the curve

$$\overline{V^2_{\text{obs.}}}/\overline{V^2_{\text{calc.}}} = e^{-2.9\times10^3 I_0^2}$$

where e is the base of Naperian logarithms, $\overline{V^2_{\text{obs.}}}$ is the observed value of mean square voltage and $\overline{V^2_{\text{calc.}}}$ is computed from the limit of the observed (voltage)^2 as I_0 approaches zero.

Acknowledgments.—The author wishes to acknowledge his indebtedness to the staff of Cruft Laboratory for assistance and advice in this work, and to Dr. Hull of the General Electric Company for the screen-grid tubes which made possible the experimental part, and in particular to Paul S. Bauer, B.C.E., for his coöperation during the latter part of the year making possible independent checks on all phases of the problem.

* NATIONAL RESEARCH FELLOW.

[1] N. H. Williams and H. B. Vincent, *Phys. Rev.*, **28**, 1250–64 (1926). This paper contains a complete introduction to the subject, together with references to other works.

[2] A. W. Hull and N. H. Williams, *Phys. Rev.*, **25**, 147 (1925).

[3] W. Schottky, *Ann. Physik*, **57**, 541–67 (1918); **68**, 157–76 (1922).

[4] See paper quoted under [1] for more complete explanation.

[5] See [2]; also J. B. Johnson, *Phys. Rev.*, **26**, 71 (1925).